

日 本 国 特 許 庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

JC997 U.S. PTO  
09/769378  
01/26/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日  
Date of Application: 2000年 1月28日

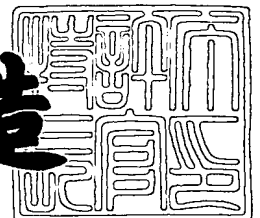
出 願 番 号  
Application Number: 特願2000-020300

出 願 人  
Applicant (s): 株式会社東芝

2000年 9月 1日

特許庁長官  
Commissioner,  
Patent Office

及川耕造



出証番号 出証特2000-3070724

【書類名】 特許願

【整理番号】 A000000104

【提出日】 平成12年 1月28日

【あて先】 特許庁長官 殿

【国際特許分類】 G06K 9/00

【発明の名称】 単語認識方法および単語認識プログラムを記憶した記憶媒体

【請求項の数】 10

【発明者】

【住所又は居所】 神奈川県川崎市幸区柳町 7 0 番地 株式会社東芝柳町工場内

【氏名】 浜村 倫行

【特許出願人】

【識別番号】 000003078

【氏名又は名称】 株式会社 東芝

【代理人】

【識別番号】 100058479

【弁理士】

【氏名又は名称】 鈴江 武彦

【電話番号】 03-3502-3181

【選任した代理人】

【識別番号】 100084618

【弁理士】

【氏名又は名称】 村松 貞男

【選任した代理人】

【識別番号】 100068814

【弁理士】

【氏名又は名称】 坪井 淳

【選任した代理人】

【識別番号】 100092196

【弁理士】

【氏名又は名称】 橋本 良郎

【選任した代理人】

【識別番号】 100091351

【弁理士】

【氏名又は名称】 河野 哲

【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 単語認識方法および単語認識プログラムを記憶した記憶媒体

【特許請求の範囲】

【請求項 1】 認識すべき単語に対応した入力文字列を各文字ごとに認識処理を行なうステップと、

あらかじめ認識すべき単語の候補が格納されている単語辞書内の各単語の各文字を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、

この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、

前記単語辞書内の各単語の各文字に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、

この演算結果に基づき前記単語の認識結果を得るステップと、

を具備したことを特徴とする単語認識方法。

【請求項 2】 認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、

この文字切出しによって文字間の特徴を抽出するステップと、

前記文字切出しによって得られる各文字に対してそれぞれ認識処理を行なうステップと、

あらかじめ認識すべき単語と単語の文字間の特徴の候補が格納されている単語辞書内の各単語の各文字および文字間の特徴を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、

この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、

前記単語辞書内の各単語の各文字および文字間の特徴に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、

この演算結果に基づき前記単語の認識結果を得るステップと、

を具備したことを特徴とする単語認識方法。

【請求項 3】 前記単語辞書内の各単語の各文字には、文字に加えて非文字の情報を含ませたことを特徴とする請求項 1 記載の単語認識方法。

【請求項 4】 前記非文字の情報を含む文字の単語の出現確率は、非文字の情報を含まない文字の単語の出現確率を基に設定されることを特徴とする請求項 3 記載の単語認識方法。

【請求項 5】 認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、

この文字切出しによって文字間の接離を考慮した複数種類の切出し結果を求めるステップと、

この求めた全ての切出し結果の各文字に対してそれぞれ認識処理を行なうステップと、

あらかじめ認識すべき単語の候補が格納されている単語辞書内の各単語の各文字を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、

この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、

前記単語辞書内の各単語の各文字に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、

この演算結果に基づき前記単語の認識結果を得るステップと、

を具備したことを特徴とする単語認識方法。

【請求項 6】 認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、

この文字切出しによって文字間の接離を考慮した複数種類の切出し結果を求めるステップと、

この求めた全ての切出し結果に対してそれぞれ文字間の特徴を抽出するステップと、

前記求めた全ての切出し結果の各文字に対してそれぞれ認識処理を行なうステップと、

あらかじめ認識すべき単語と単語の文字間の特徴の候補が格納されている単語

辞書内の各単語の各文字および文字間の特徴を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、

この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、

前記単語辞書内の各単語の各文字および文字間の特徴に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、

この演算結果に基づき前記単語の認識結果を得るステップと、

を具備したことを特徴とする単語認識方法。

【請求項 7】 コンピュータに単語認識処理を行なわせるための単語認識プログラムを記憶したコンピュータ読取可能な記憶媒体であって、

前記単語認識プログラムは、

認識すべき単語に対応した入力文字列を各文字ごとに認識処理を行なうステップと、

あらかじめ認識すべき単語の候補が格納されている単語辞書内の各単語の各文字を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、

この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、

前記単語辞書内の各単語の各文字に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、

この演算結果に基づき前記単語の認識結果を得るステップと、

を有することを特徴とする単語認識プログラムを記憶した記憶媒体。

【請求項 8】 コンピュータに単語認識処理を行なわせるための単語認識プログラムを記憶したコンピュータ読取可能な記憶媒体であって、

前記単語認識プログラムは、

認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、

この文字切出しによって文字間の特徴を抽出するステップと、

前記文字切出しによって得られる各文字に対してそれぞれ認識処理を行なうステップと、

あらかじめ認識すべき単語と単語の文字間の特徴の候補が格納されている単語辞書内の各単語の各文字および文字間の特徴を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、

この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、

前記単語辞書内の各単語の各文字および文字間の特徴に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、

この演算結果に基づき前記単語の認識結果を得るステップと、

を有することを特徴とする単語認識プログラムを記憶した記憶媒体。

【請求項 9】 コンピュータに単語認識処理を行なわせるための単語認識プログラムを記憶したコンピュータ読取可能な記憶媒体であって、

前記単語認識プログラムは、

認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、

この文字切出しによって文字間の接離を考慮した複数種類の切出し結果を求めるステップと、

この求めた全ての切出し結果の各文字に対してそれぞれ認識処理を行なうステップと、

あらかじめ認識すべき単語の候補が格納されている単語辞書内の各単語の各文字を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、

この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、

前記単語辞書内の各単語の各文字に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、

この演算結果に基づき前記単語の認識結果を得るステップと、

を有することを特徴とする単語認識プログラムを記憶した記憶媒体。

【請求項 10】 コンピュータに単語認識処理を行なわせるための単語認識プログラムを記憶したコンピュータ読取可能な記憶媒体であって、

前記単語認識プログラムは、

認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、  
 この文字切出しによって文字間の接離を考慮した複数種類の切出し結果を求めるステップと、  
 この求めた全ての切出し結果に対してそれぞれ文字間の特徴を抽出するステップと、  
 前記求めた全ての切出し結果の各文字に対してそれぞれ認識処理を行なうステップと、  
 あらかじめ認識すべき単語と単語の文字間の特徴の候補が格納されている単語辞書内の各単語の各文字および文字間の特徴を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、  
 この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、  
 前記単語辞書内の各単語の各文字および文字間の特徴に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、  
 この演算結果に基づき前記単語の認識結果を得るステップと、  
 を有することを特徴とする単語認識プログラムを記憶した記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、たとえば、被読取物上に記載された複数の文字からなる単語を光学的に読取る光学的文字読取装置などにおいて単語認識を行なう単語認識方法、および、その単語認識処理を行なわせるための単語認識プログラムを記憶した記憶媒体に関する。

【0002】

【従来の技術】

一般に、たとえば、光学的文字読取装置において、被読取物上に記載された文字を読取る場合、個々の文字認識精度が低くても、単語の知識を用いることで精度良く読取ることができる。従来、その実施方法として種々の方法が提案されている。



【0003】

たとえば、特開平10-177624号公報に開示されるものは、文字認識の結果として距離（その値が小さいほど認識結果が信頼できるもの）を用い、その距離の和をもって単語の評価値としている。

【0004】

また、たとえば、特開平8-167008号公報に開示されるものは、文字認識の段階で各文字の候補を絞り、絞られた候補と各単語との照合を行ない、一致する文字数をもって単語の評価値としている。

【0005】

さらに、たとえば、電子通信学会論文誌Vol.52-C, No.6, June 1969, P.305-312に開示されるものは、単語の評価値として事後確率を用いている。

【0006】

【発明が解決しようとする課題】

しかしながら、従来の方法のいずれも、単語の文字数が一定であるときにのみ意味のあるものであり、文字数が一定でない場合には用いることができないか、用いても不具合が発生する。すなわち、

特開平10-177624号公報に開示されるものは、文字数が少ないほど評価値は小さくなるため、文字数の少ない単語が選ばれやすくなっている。

【0007】

また、特開平8-167008号公報、電子通信学会論文誌に開示されるものは、文字数が一定であることが前提となっており、文字数が一定でないときには用いることができない。

【0008】

さらに、従来の単語認識の評価関数においては、単語切出しの曖昧さ、文字接触、ノイズ混入、文字切出しの曖昧さは考慮されていない。

【0009】

そこで、本発明は、単語の文字数が一定でない場合においても精度良く単語認識が行なえる単語認識方法および単語認識プログラムを記憶した記憶媒体を提供することを目的とする。

【0010】

また、本発明は、単語切出しが確実にない場合にも精度良く単語認識が行なえる単語認識方法および単語認識プログラムを記憶した記憶媒体を提供することを目的とする。

【0011】

また、本発明は、文字接触やノイズ混入がある場合にも精度良く単語認識が行なえる単語認識方法および単語認識プログラムを記憶した記憶媒体を提供することを目的とする。

【0012】

さらに、本発明は、文字切出しが確実にない場合にも精度良く単語認識が行なえる単語認識方法および単語認識プログラムを記憶した記憶媒体を提供することを目的とする。

【0013】

【課題を解決するための手段】

本発明の単語認識方法は、認識すべき単語に対応した入力文字列を各文字ごとに認識処理を行なうステップと、あらかじめ認識すべき単語の候補が格納されている単語辞書内の各単語の各文字を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、前記単語辞書内の各単語の各文字に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、この演算結果に基づき前記単語の認識結果を得るステップとを具備している。

【0014】

また、本発明の単語認識方法は、認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、この文字切出しによって文字間の特徴を抽出するステップと、前記文字切出しによって得られる各文字に対してそれぞれ認識処理を行なうステップと、あらかじめ認識すべき単語と単語の文字間の特徴の候補が格納されている単語辞書内の各単語の各文字および文字間の特徴を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、この求めた確

率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、前記単語辞書内の各単語の各文字および文字間の特徴に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、この演算結果に基づき前記単語の認識結果を得るステップとを具備している。

## 【 0 0 1 5 】

また、本発明の単語認識方法は、認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、この文字切出しによって文字間の接離を考慮した複数種類の切出し結果を求めるステップと、この求めた全ての切出し結果の各文字に対してそれぞれ認識処理を行なうステップと、あらかじめ認識すべき単語の候補が格納されている単語辞書内の各単語の各文字を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、前記単語辞書内の各単語の各文字に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、この演算結果に基づき前記単語の認識結果を得るステップとを具備している。

## 【 0 0 1 6 】

また、本発明の単語認識方法は、認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、この文字切出しによって文字間の接離を考慮した複数種類の切出し結果を求めるステップと、この求めた全ての切出し結果に対してそれぞれ文字間の特徴を抽出するステップと、前記求めた全ての切出し結果の各文字に対してそれぞれ認識処理を行なうステップと、あらかじめ認識すべき単語と単語の文字間の特徴の候補が格納されている単語辞書内の各単語の各文字および文字間の特徴を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、前記単語辞書内の各単語の各文字および文字間の特徴に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、この演算結果に基づき前記単語の認識結果を得るステップとを具備している。

## 【 0 0 1 7 】

また、本発明の記憶媒体は、コンピュータに単語認識処理を行なわせるための単語認識プログラムを記憶したコンピュータ読取可能な記憶媒体であって、前記単語認識プログラムは、認識すべき単語に対応した入力文字列を各文字ごとに認識処理を行なうステップと、あらかじめ認識すべき単語の候補が格納されている単語辞書内の各単語の各文字を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、前記単語辞書内の各単語の各文字に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、この演算結果に基づき前記単語の認識結果を得るステップとを有する。

## 【 0 0 1 8 】

また、本発明の記憶媒体は、コンピュータに単語認識処理を行なわせるための単語認識プログラムを記憶したコンピュータ読取可能な記憶媒体であって、前記単語認識プログラムは、認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、この文字切出しによって文字間の特徴を抽出するステップと、前記文字切出しによって得られる各文字に対してそれぞれ認識処理を行なうステップと、あらかじめ認識すべき単語と単語の文字間の特徴の候補が格納されている単語辞書内の各単語の各文字および文字間の特徴を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、前記単語辞書内の各単語の各文字および文字間の特徴に対してそれぞれ求めた前記各演算結果の間で所定の演算を行なうステップと、この演算結果に基づき前記単語の認識結果を得るステップとを有する。

## 【 0 0 1 9 】

また、本発明の記憶媒体は、コンピュータに単語認識処理を行なわせるための単語認識プログラムを記憶したコンピュータ読取可能な記憶媒体であって、前記単語認識プログラムは、認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、この文字切出しによって文字間の接離を考慮した複数種類の切出し結果を求めるステップと、この求めた全ての切出し結果の各文字に対してそ

れぞれ認識処理を行なうステップと、あらかじめ認識すべき単語の候補が格納されている単語辞書内の各単語の各文字を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、前記単語辞書内の各単語の各文字に対してそれぞれ求めた前記各演算結果の間に所定の演算を行なうステップと、この演算結果に基づき前記単語の認識結果を得るステップとを有する。

#### 【0020】

さらに、本発明の記憶媒体は、コンピュータに単語認識処理を行なわせるための単語認識プログラムを記憶したコンピュータ読取可能な記憶媒体であって、前記単語認識プログラムは、認識すべき単語に対応した入力文字列を各文字ごとに切出すステップと、この文字切出しによって文字間の接離を考慮した複数種類の切出し結果を求めるステップと、この求めた全ての切出し結果に対してそれぞれ文字間の特徴を抽出するステップと、前記求めた全ての切出し結果の各文字に対してそれぞれ認識処理を行なうステップと、あらかじめ認識すべき単語と単語の文字間の特徴の候補が格納されている単語辞書内の各単語の各文字および文字間の特徴を条件として文字認識結果として得られた特徴が出現する確率を求めるステップと、この求めた確率と文字認識結果として得られた特徴が出現する確率との間で所定の演算を行なうステップと、前記単語辞書内の各単語の各文字および文字間の特徴に対してそれぞれ求めた前記各演算結果の間に所定の演算を行なうステップと、この演算結果に基づき前記単語の認識結果を得るステップとを有する。

#### 【0021】

本発明によれば、文字認識結果を用いた単語認識において、単語の文字数が一定とは限らない場合にも用いることのできる事後確率を基にした評価関数を用いることにより、単語の文字数が一定でない場合においても精度良く単語認識が行なえる。

#### 【0022】

また、本発明によれば、文字認識結果を用いた単語認識において、少なくとも

単語切出しの曖昧さをも考慮した事後確率を基にした評価関数を用いることにより、単語切出しが確実でない場合にも精度良く単語認識が行なえる。

【 0 0 2 3 】

また、本発明によれば、文字認識結果を用いた単語認識において、少なくとも文字接触をも考慮した事後確率を基にした評価関数を用いることにより、文字接触がある場合にも精度良く単語認識が行なえる。

【 0 0 2 4 】

また、本発明によれば、文字認識結果を用いた単語認識において、少なくともノイズ混入をも考慮した事後確率を基にした評価関数を用いることにより、ノイズ混入がある場合にも精度良く単語認識が行なえる。

【 0 0 2 5 】

さらに、本発明によれば、文字認識結果を用いた単語認識において、少なくとも文字切出しの曖昧さをも考慮した事後確率を基にした評価関数を用いることにより、文字切出しが確実でない場合にも精度良く単語認識が行なえる。

【 0 0 2 6 】

【発明の実施の形態】

以下、本発明の実施の形態について図面を参照して説明する。

【 0 0 2 7 】

図 1 は、本発明の実施の形態に係る単語認識方法を実現するための単語認識システムの構成を概略的に示すものである。

【 0 0 2 8 】

図 1 において、この単語認識システムは、CPU（セントラル・プロセッシング・ユニット）1、入力装置2、画像入力手段としてのスキャナ3、表示装置4、記憶手段としての第1メモリ5、記憶手段としての第2メモリ6、および、読取装置7などによって構成されている。

【 0 0 2 9 】

CPU 1 は、第 2 メモリ 6 に格納されたオペレーティングシステムプログラム、および、第 2 メモリ 6 に格納されたアプリケーションプログラム（単語認識プログラムなど）を実行することにより、後で詳細を説明するような単語認識処理

などを行なうものである。

【 0 0 3 0 】

入力装置 2 は、たとえば、キーボードやマウスなどからなり、利用者が各種操作を行なったり、各種データを入力するために使用される。

【 0 0 3 1 】

スキャナ 3 は、被読取物上に記載された単語の各文字を光学的な走査により読取って入力するものである。

【 0 0 3 2 】

表示装置 4 は、たとえば、ディスプレイ装置やプリンタなどからなり、各種データを出力するものである。

【 0 0 3 3 】

第 1 メモリ 5 は、たとえば、RAM（ランダム・アクセス・メモリ）により構成されていて、CPU 1 の作業用メモリとして使用され、処理中の各種データなどを一時記憶するものである。

【 0 0 3 4 】

第 2 メモリ 6 は、たとえば、ハードディスク装置により構成されていて、CPU 1 を動作させるための各種プログラムなどを記憶するものである。第 2 メモリ 6 には、入力装置 2、スキャナ 3、表示装置 4、第 1 メモリ 5、第 2 メモリ 6、読取装置 7 などを動作させるためのオペレーティングシステムプログラム、単語認識プログラムおよび単語認識用の単語辞書などが記憶されている。

【 0 0 3 5 】

読取装置 7 は、たとえば、CD-ROM ドライブ装置などからなり、記憶媒体としての CD-ROM 8 に格納（記憶）されている単語認識プログラムおよび単語認識用の単語辞書などを読取るものである。読取装置 7 により読取られた単語認識プログラムおよび単語認識用の単語辞書は、第 2 メモリ 6 に格納（記憶）される。なお、単語辞書には、あらかじめ認識すべき単語の候補が複数格納されている。

【 0 0 3 6 】

本発明の実施の形態に係る単語認識方法は、このようなシステム構成において

実現されるものであるが、まず概要について説明する。

【0037】

1. 概要

たとえば、光学的文字読取装置による文字読取りを考える。文字認識の性能が高く、ほとんど読み間違えないのならば問題はないが、たとえば、手書文字の認識ではそこまでの性能を得るのが難しい。そこで、単語の知識を用いることで認識精度を高める。具体的には、単語の辞書の中から正しいと考えられる単語を選び出す。そのために、各単語ごとに何らかの評価値を計算し、最も高い（低い）評価値をとった単語を認識結果とすることになる。評価関数として、従来は前述したような各種のものが考えられているが、前述したような各種問題がある。

【0038】

そこで、本実施の形態では、評価関数として前述した各種問題を考慮した事後確率を用いる。これにより、文字数の違い、単語切出しの曖昧さ、文字接触、ノイズ、文字分離までを全て確率の計算により1つの評価関数に自然に組込むことができる。

【0039】

次に、本発明で用いるノイズ推定の一般論について説明する。

【0040】

2. ベイズ推定の一般論

入力パターン（入力文字列）を $x$ とする。認識処理とは、 $x$ に対しある処理を行ない、分類結果が得られるものである。これを分解して考えてみると、以下の2つのプロセスに分けることができる。

【0041】

- (1)  $x$ について何らかの特徴量を得る特徴抽出処理 $R$ をかけることで、「特徴」 $r (= R(x))$ を得る。
- (2) 「特徴」 $r$ を見てそこに何らかの評価法を用い、分類結果 $k_i$ を得る。

【0042】

分類結果 $k_i$ が「認識結果」である。単語認識においては、「特徴」の1つとして文字認識の「認識結果」が使われるので注意が必要である。今後、用語とし



ての「特徴」と「認識結果」を区別して用いる。

【0043】

ベイズ推定は、2番目のプロセスにおける評価法として用いられる。事後確率  $P(k_i | r)$  が最大となるカテゴリ  $k_i$  を認識結果とする。事後確率  $P(k_i | r)$  を直接計算するのが困難、もしくは不可能である場合、ベイズの定理、

【0044】

【数1】

$$P(k_i | r) = \frac{P(r | k_i)P(k_i)}{P(r)} \quad \text{式(1)}$$

【0045】

を用いて間接的に計算する。分母の  $P(r)$  は  $i$  によらない定数であるため、分子の  $P(r | k_i)P(k_i)$  を計算することで、事後確率  $P(k_i | r)$  の大小を評価することができる。

【0046】

次に、以後の説明の理解を容易にするため、文字数が一定であった場合の単語認識におけるベイズ推定について説明する。この場合のベイズ推定は、英語などの、単語を分けて書く言語において有用である。

【0047】

### 3. 文字数一定の場合のベイズ推定

#### 3. 1. 定式化

文字の切出し、単語の切出しが完全に成功しており、また文字間のノイズの混入もなく、文字数が確定していると仮定する。次のように定義する。

【0048】

## 【数2】

- ・文字数  $L$
- ・カテゴリ集合  $K = \{k_i\}$
- $k_i = \hat{w}_i, \hat{w}_i \in \hat{W}, \hat{W} : \text{文字数 } L \text{ のワードの集合}$
- ・  $\hat{w}_i = (\hat{w}_{i1}, \hat{w}_{i2}, \dots, \hat{w}_{iL})$
- $\hat{w}_{ij} : \hat{w}_i \text{ の } j \text{ 番目の文字 } \hat{w}_{ij} \in C, C : \text{文字の集合}$
- ・特徴  $r = (r_1, r_2, r_3, \dots, r_L)$

$r_i : i \text{ 文字目の文字特徴 (＝文字認識結果)}$

(例. 1 位候補、3 位までの候補、類似度一定値以上の候補、  
1, 2 位候補とその類似度、等)

なお、以後の説明文中において、 $\hat{w}_i$  を  $w_i$  と表記することもある。

【0049】

このとき、ベイズ推定によって書かれている単語を推定することを考える。

【0050】

## 【数3】

$$P(k_i | r) = \frac{P(r | k_i) P(k_i)}{P(r)} \quad \text{式 (2)}$$

【0051】

$P(r | k_i)$  は次のように表わされる。

【0052】

## 【数4】

$$P(r | k_i) = P(r_1 | \hat{w}_{i1}) P(r_2 | \hat{w}_{i2}) \cdots P(r_L | \hat{w}_{iL}) = \prod_{j=1}^L P(r_j | \hat{w}_{ij}) \quad \text{式 (3)}$$

【0053】

$P(k_i)$  は統計的に求まっているものとする。たとえば、郵便物の宛名読取りの場合、宛先の統計だけでなく、書状内での位置、行内での位置などに依存することも考えられる。

【0054】

なお、 $P(r | k_i)$  は積で表わされているが、これに限らず、たとえば、対

数を取り、加算に変換することも可能である。以下の説明においても同じことが言える。

【0055】

### 3. 2. 実用に向けた近似

特徴  $r_i$  として何を用いるかにより、認識の性能に大きな違いがでる。

【0056】

#### 3. 2. 1. 1位候補を用いた場合

文字特徴  $r_i$  として「1位候補の文字」を用いることを考えてみる。次のように定義する。

【0057】

・文字集合  $C = \{c_i\}$

例  $c_i$  : 数字、 $c_i$  : アルファベット大文字小文字など

・文字特徴集合  $E = \{e_i\}$

$e_i = (1\text{位候補が } c_i)$

・  $r_i \in E$

たとえば、文字集合  $C$  として、「アルファベット大文字小文字+数字」を想定すると、特徴  $e_i$  の種類も文字  $c_i$  の種類も  $n(C) = n(E) = 62$  通りであるため、 $(e_i, c_j)$  の組合わせは「62」の二乗通りであり、「62」の二乗通りの  $P(e_i | c_j)$  をあらかじめ用意しておくことで、前記式(3)は計算される。具体的には、たとえば、 $P(e_i | "A")$  を求めるには、“A”のサンプルをたくさん特徴抽出処理  $R$  に流し、各特徴  $e_i$  の出現頻度を調べればよい。

【0058】

#### 3. 2. 2. 近似

ここで、次のような近似を用いることも考えられる。

【0059】

【数5】

$$\forall i, P(e_i | c_i) = p$$

式(4)

【0060】

【数6】

$$\forall i \neq j, P(e_j | c_i) = q \quad \text{式(5)}$$

【0061】

上記式(4)、式(5)は、どの文字  $c_i$  でも、1位候補がその文字自体になる確率は等しく  $p$  であり、それ以外の文字になる確率は等しく  $q$  である、という近似である。このとき、

$$p + \{n(E) - 1\} q = 1$$

式(6)

が成り立っている。

【0062】

この近似は、1位候補を並べた文字列を仮の認識結果とし、その文字列と各ワード  $w_a$  が何文字一致しているかを見るマッチングと対応している。a文字一致したとき、

【0063】

【数7】

$$P(r | \hat{w}_i) = p^a q^{L-a} \quad \text{式(7)}$$

【0064】

と表わされ、簡便である。

【0065】

### 3. 3. 具体例

たとえば、英語で書かれた郵便物の宛名読取りにおいて、都市名を読むことを考える。図2は、都市名が書かれていると考えられる部分であり、4つの文字であることがわかる。図3は、単語辞書の内容であり、4つの文字からなる都市名(単語)の候補が格納されている。

【0066】

図2の各文字を文字認識にかけ、その文字認識結果を用いて図3の各都市名の事後確率を計算する。計算に用いる特徴(=文字認識結果)は様々であるが、こ

ここでは1位候補の文字を用いた例を示す。図4は、各文字の認識結果を示している。たとえば、一番左の文字は、認識結果の1位が「H」であったことを示す。都市名が「MAIR」であったときに、図4の認識結果が生じる確率 $P(r | k_i)$ は、前記式(3)より、

【0067】

【数8】

$$P(r | k_1) = P("H" | "M")P("A" | "A")P("I" | "I")P("A" | "R")$$

式(8)

【0068】

である。3. 2. 1節で述べたように、あらかじめ右辺の各項の値は求めておく。もしくは、3. 2. 2節で述べた近似を用いれば、たとえば、 $p = 0.5$ 、 $n(E) = 26$ のときは、 $q = 0.02$ であるから、

$$P(r | k_1) = q \cdot p \cdot p \cdot q = 0.0001$$

式(9)

同様にして

$$P(r | k_2) = q \cdot q \cdot q \cdot p = 0.000004$$

$$P(r | k_3) = q \cdot q \cdot q \cdot q = 0.00000016$$

$$P(r | k_4) = p \cdot p \cdot q \cdot p = 0.0025$$

$$P(r | k_5) = p \cdot q \cdot q \cdot q = 0.000004$$

式(10)

$P(k_1) \sim P(k_5)$ を等しいと見なせば、前記式(2)より事後確率 $P(k_i | r)$ の大小は $P(r | k_i)$ と同じである。したがって、上記式(9)、式(10)の大小を比較すればよく、最も大きいのは $P(r | k_4)$ であるので、図2に書かれている都市名は「HAMA」とであると推定される。

【0069】

3. 2. 2節の近似を用いた結果、図4と各都市名の間で一致している文字数が多い都市名が選ばれることになっている。3. 2. 2節の近似を用いず、3. 2. 1節で述べたように、あらかじめ各 $P(e_i | c_j)$ を求めて、それを計算に用いた場合、一致した文字数が多ければよいとは限らない。

【0070】

たとえば、前記式(8)の第1項は、「H」と「M」は形状が似ていることから比較的大きい値であり、

【0071】

【数9】

$$P("M"|"M") = 0.32, P("H"|"M") = 0.2,$$

$$P("H"|"H") = 0.32, P("M"|"H") = 0.2$$

【0072】

とする。また、第4項も同様であり、

【0073】

【数10】

$$P("R"|"R") = 0.42, P("A"|"R") = 0.1,$$

$$P("A"|"A") = 0.42, P("R"|"A") = 0.1$$

【0074】

であるとする。他の文字に関しては、3. 2. 2節の近似が成り立つとする。このとき、

【0075】

【数11】

$$P(r|k_1) = P("H"|"M") \cdot p \cdot p \cdot P("A"|"R") = 0.005$$

$$P(r|k_2) = q \cdot q \cdot q \cdot P("A"|"A") = 0.00000336$$

$$P(r|k_3) = q \cdot q \cdot q \cdot q = 0.00000016$$

$$P(r|k_4) = P("H"|"H") \cdot P("A"|"A") \cdot q \cdot P("A"|"A") \approx 0.0011$$

$$P(r|k_5) = P("H"|"H") \cdot q \cdot q \cdot q = 0.00000256$$

式(11)

【0076】

であり、 $P(r|k_1)$ が最も大きな値となり、図2に書かれていると推定される都市名は「MAIR」となる。

【0077】

次に、本発明の第1の実施の形態としての、文字数が一定でない場合の単語認識におけるベイズ推定について説明する。この場合のベイズ推定は、日本語などの、単語を分けて書かない言語において有用である。また、単語を分けて書く言語においても、単語辞書に複数単語からなる文字列がある場合は有用である。

【0078】

#### 4. 文字数が一定でない場合のベイズ推定

実際には、複数単語の文字列がカテゴリに含まれる場合（例：NORTH YORK）もあるが、3節で述べた方法では1単語の文字列と2単語の文字列での比較はできない。また、単語を分けて書かない言語（日本語など）では文字数が一定とならず、3節の方法は使えない。そこで、本節では文字数が一定とは限らない場合に対応した単語認識方法を説明する。

【0079】

##### 4. 1. 定式化

入力パターン $x$ を、1単語ではなく、複数単語とし、3節と同様にベイズ推定を行なう。この場合、3節の定義を次のように追加変更する。

【0080】

変更

- ・入力パターン $x$ を複数単語とする。
- ・ $L$ ：入力パターン $x$ 内の全文字数
- ・カテゴリ集合 $K = \{k_i\}$

【0081】

【数12】

$$k_i = (\hat{w}'_j, h)$$

$\hat{w}'_j \in \hat{w}'$ ,  $\hat{w}'$ ：入力 $x$ に当てはめることのできる文字数、単語数を持つ

文字列の集合

$h$ ：入力 $x$ 内での文字列 $\hat{w}'_j$ の位置。入力 $x$ の先頭から $(h + 1)$ 文字目から文字列 $\hat{w}'_j$ が開始する。

なお、以後の説明文中において、 $\hat{w}'_j$ を $w_b$ とすることもある。

【0082】

追加

【0083】

【数13】

$$\begin{aligned} \cdot \hat{w}'_j &= (\hat{w}'_{j1}, \hat{w}'_{j2}, \dots, \hat{w}'_{jL_j}) \\ L_j &: \text{文字列 } \hat{w}'_j \text{ の全文字数} \\ \hat{w}'_{jk} &: \hat{w}'_j \text{ の } k \text{ 番目の文字 } \hat{w}'_{jk} \in C \end{aligned}$$

【0084】

このとき、ベイズ推定を用いると、事後確率  $P(k_i | r)$  は、前記式(2)と同じである。

【0085】

【数14】

$$P(k_i | r) = \frac{P(r | k_i) P(k_i)}{P(r)} \quad \text{式 (12)}$$

【0086】

$P(r | k_i)$  は次のように表わされる。

【0087】

【数15】

$$\begin{aligned} P(r | K_i) &= P(r_1, r_2, \dots, r_h | k_i) \\ &\cdot P(r_{h+1} | \hat{w}'_{j1}) P(r_{h+2} | \hat{w}'_{j2}) \dots P(r_{h+L_j} | \hat{w}'_{jL_j}) \\ &\cdot P(r_{h+L_j+1}, r_{h+L_j+2}, \dots, r_L | k_i) \\ &= P(r_1, r_2, \dots, r_h | k_i) \left\{ \prod_{k=1}^{L_j} P(r_{h+k} | \hat{w}'_{jk}) \right\} \\ &\cdot P(r_{h+L_j+1}, r_{h+L_j+2}, \dots, r_L | k_i) \end{aligned}$$

式 (13)

【0088】

$P(k_i)$  は、3節と同様に求まっているものとする。 $n(K)$  は、3節よりも増えるため、単純には  $P(k_i)$  は3節よりも小さな値となることに注意する



【0089】

## 4. 2. 実用に向けた近似

## 4. 2. 1. 文字列のない部分に対する近似と文字数正規化

上記式 (13) の第1項を次のように近似する。

【0090】

【数16】

$$\begin{aligned} P(x_1, x_2, \dots, x_h | k_i) &\approx P(x_1, x_2, \dots, x_h) \\ &\approx P(x_1)P(x_2) \cdots P(x_h) \end{aligned} \quad \text{式 (14)}$$

【0091】

1行目の近似は、「入力パターン  $x$  の全文字のうち文字列  $wb$  が当てはまっていない部分への、 $wb$  による影響を無視する」というものである。2行目は、「各  $r_k$  が独立である」というものである。実際にはそうではなく、粗い近似ではあるが、非常に有効である。

【0092】

同様に、前記式 (13) の第3項も近似すると、前記式 (13) は次のようになる。

【0093】

【数17】

$$P(x | k_i) = \prod_{k=1}^{L_j} P(x_{h+k} | \hat{w}_{jk}) \prod_{\substack{1 \leq k \leq h \\ h+L_j+1 \leq k \leq L}} P(x_k) \quad \text{式 (15)}$$

【0094】

ここで、 $P(k_i | r) / P(k_i)$  という値を考える。これは、 $k_i$  である確率が、特徴  $r$  を知ることでどれほど増減したかを示す値である。

【0095】

【数 18】

$$\begin{aligned}
 \frac{P(k_i | r)}{P(k_i)} &= \frac{P(r | k_i)}{P(r)} \\
 &\approx \frac{\prod_{k=1}^{L_j} P(x_{h+k} | \hat{w}'_{jk}) \prod_{\substack{1 \leq k \leq h \\ h+L_j+1 \leq k \leq L}} P(x_k)}{\prod_{k=1}^L P(x_k)} \quad \text{式 (16)} \\
 &= \prod_{k=1}^{L_j} \frac{P(x_{h+k} | \hat{w}'_{jk})}{P(x_{h+k})}
 \end{aligned}$$

【0096】

上記式 (16) の 2 行目の分母で用いている近似は、前記式 (14) と同様である。

【0097】

この結果は非常に重要である。上記式 (16) の右辺には、全文字のうち文字列  $w_b$  が当てはまっていない部分に関する記述はない。すなわち、上記式 (16) は、入力パターン  $x$  が何であるかに無関係である。このことから、文字列  $w_b$  の位置、長さは気にせずに上記式 (16) を計算し、 $P(k_i)$  をかけることで、 $P(k_i | r)$  を計算できることがわかる。

【0098】

前記式 (16) の分子を見ると、前記式 (3)、つまり文字数を一定とした場合の  $P(r | k_i)$  と同じ式になっている。これは、前記式 (16) は分母によって文字数の正規化を行なっているといえる。

【0099】

#### 4. 2. 2. 1 位候補を用いた場合

ここで、特徴として 3. 2. 1 節のように「1 位候補の文字」を用いたとする。すると、 $P(r_k)$  の近似として次のものが考えられる。

【0100】

【数19】

$$P(r_k) = \frac{1}{n(E)} \quad \text{式 (17)}$$

【0101】

実際には、各文字の生起確率を考慮する必要があり、それを無視したものである。このとき、3. 2. 2節の近似も用いて前記式(16)を近似すると、次の式となる。

【0102】

【数20】

$$\frac{P(k_i|r)}{P(k_i)} = p^a q^{L_j - a} n(E)^{L_j} \quad \text{式 (18)}$$

【0103】

$n(E)^{L_j}$ により正規化がなされている。

【0104】

#### 4. 3. 具体例

たとえば、郵便物の宛名読取りにおいて、都市名を読むことを考える。

【0105】

・単語を分けて書く言語（英語など）であり、複数単語の都市名が存在するとき

・単語を分けて書かない言語（日本語など）のとき

以上のとき、各候補の文字数は一定にはならない。例として、英語の宛名読取りを上げる。図5は、都市名が書かれていると考えられる部分であり、2文字単語の後、3文字単語が存在することがわかる。図6は、単語辞書の内容であり、図5に当てはまる文字数、単語数を持つ都市名が全て格納されている。

【0106】

図5の各文字を文字認識にかけ、その文字認識結果を用いて図6の各都市名の事後確率を計算する。計算に用いる特徴（＝文字認識結果）は様々であるが、ここでは1位候補の文字を用いた例を示す。図7は、各文字の認識結果を示している。たとえば、一番左の文字は、認識結果の1位が「S」であったことを示す。

都市名が「COH」であり、右の3文字に「COH」が書かれているという事象が図7の認識結果から生じる事後確率 $P(k_i | r)$ は、4.2.1節で述べた近似を用いると、前記式(16)により、

【0107】

【数21】

$$\frac{P(k_1|r)}{P(k_1)} \approx \frac{P("C"|"C")}{P("C")} \frac{P("T"|"O")}{P("T")} \frac{P("H"|"H")}{P("H")} \quad \text{式(19)}$$

【0108】

から計算できる。さらに、3.2.2節、4.2.2節で述べた近似を用いれば、たとえば、 $p=0.5$ 、 $n(E)=26$ のときは、 $q=0.02$ であるから、

【0109】

【数22】

$$\frac{P(k_1|r)}{P(k_1)} \approx p \cdot q \cdot p \cdot n(E)^3 = 87.88 \quad \text{式(20)}$$

【0110】

同様にして

【0111】

【数23】

$$\begin{aligned} \frac{P(k_2|r)}{P(k_2)} &\approx q \cdot q \cdot q \cdot p \cdot p \cdot n(E)^5 \approx 23.76 \\ \frac{P(k_3|r)}{P(k_3)} &\approx q \cdot p \cdot p \cdot n(E)^3 = 87.88 \\ \frac{P(k_4|r)}{P(k_4)} &\approx p \cdot p \cdot n(E)^2 = 169 \end{aligned} \quad \text{式(21)}$$

$$\frac{P(k_5|r)}{P(k_5)} \approx p \cdot q \cdot q \cdot q \cdot q \cdot n(E)^5 \approx 0.95$$

【0112】

ただし、 $k_3$ は右3文字が「OTH」であり、 $k_4$ は左2文字が「SK」であるとしたものである。

【0113】

$P(k_1) \sim P(k_5)$  を等しいと見なせば、事後確率  $P(k_i | r)$  の大小は前記式(20)、式(21)の大小を比較すればよく、最も大きいのは  $P(k_4 | r)$  であるので、図5に書かれている都市名は「SK」とであると推定される。

【0114】

3. 2. 2節の近似を用いず、3. 2. 1節で述べたように、あらかじめ各  $P(e_i | c_j)$  を求めて、それを計算に用いた場合の例もあげておく。

【0115】

「C」と「L」、「T」と「H」、「H」と「N」の形状が似ていることから

【0116】

【数2.4】

$$\begin{aligned} P("C"|"C") &= P("L"|"L") = P("T"|"T") = P("I"|"I") \\ &= P("H"|"H") = P("N"|"N") = 0.4 \\ P("C"|"L") &= P("L"|"C") = P("T"|"I") = P("I"|"T") \\ &= P("N"|"H") = P("H"|"N") = 0.12 \end{aligned}$$

【0117】

であるとし、他の文字に関しては、3. 2. 2節の近似が成り立つとする。このとき、

【0118】

【数25】

$$\frac{P(k_1|r)}{P(k_1)} \approx P("C"|"C") \cdot q \cdot P("H"|"H") \cdot n(E)^3 \approx 56.24$$

$$\frac{P(k_2|r)}{P(k_2)} \approx q \cdot q \cdot q \cdot P("T"|"T") \cdot P("H"|"H") \cdot n(E)^5 \approx 15.21$$

$$\frac{P(k_3|r)}{P(k_3)} \approx q \cdot P("T"|"T") \cdot P("H"|"H") \cdot n(E)^3 \approx 56.24$$

$$\frac{P(k_4|r)}{P(k_4)} \approx p \cdot p \cdot n(E)^2 \approx 169$$

$$\begin{aligned} \frac{P(k_5|r)}{P(k_5)} &\approx p \cdot q \cdot P("C"|"L") \cdot P("T"|"I") \\ &\quad \cdot P("H"|"N") \cdot n(E)^5 \approx 205.3 \end{aligned}$$

式 (22)

【0119】

であり、 $P(k_5 | r) / P(k_5)$  が最も大きな値となり、図5に書かれていると推定される都市名は「ST LIN」となる。

【0120】

このように、第1の実施の形態では、認識すべき単語に対応した入力文字列を各文字ごとに認識処理を行ない、あらかじめ認識すべき単語の候補が格納されている単語辞書内の各単語の各文字を条件として文字認識結果として得られた特徴が出現する確率を求め、この求めた確率を文字認識結果として得られた特徴が出現する確率で除算し、単語辞書内の各単語の各文字に対してそれぞれ求めた上記各除算結果を全ての文字について乗算し、この乗算結果に基づき上記単語の認識結果を得るものである。

【0121】

すなわち、文字認識結果を用いた単語認識において、単語の文字数が一定とは限らない場合にも用いることのできる、事後確率を基にした評価関数を用いることにより、単語の文字数が一定でない場合においても精度良く単語認識が行なえる。

## 【0 1 2 2】

次に、本発明の第2の実施の形態としての、単語切出しが曖昧であるとき、その曖昧さまで事後確率の計算に含めたベイズ推定について説明する。この場合のベイズ推定は、単語切れ目の誤検出が無視できないときに有用となる。

## 【0 1 2 3】

## 5. 単語切出しの統合

単語を分けて書く言語（英語など）においては、4節までで述べた方法では、単語が必ず正しく切出されているという前提があり、これが満たされず文字数に変化があると対応できない。そこで、本節では、単語切出しの結果を絶対とせず、確率として扱うことで、単語切出しの曖昧さを単語認識におけるベイズ推定に統合する。4節との大きな違いは、単語切出しの結果として得られる文字間の特徴を考慮していることである。

## 【0 1 2 4】

## 5. 1. 定式化

ここでも文字の切出しは完全に成功しており、ノイズの混入もないとする。4節の定義を基に次のように追加変更する。

## 【0 1 2 5】

変更

- ・入力パターン  $x$  を行とする。
- ・  $L$  : 入力行  $x$  内の全文字数
- ・カテゴリ集合  $K = \{k_i\}$

## 【0 1 2 6】

## 【数 2 6】

$k_i = (\tilde{w}_j, h)$ ,  $\tilde{w}_j \in \tilde{W}$ ,  $\tilde{W}$  : 文字列候補全ての集合（文字数の制限なし）

$h$  : 入力行  $x$  内の文字列  $\tilde{w}_j$  の位置。入力パターン  $x$  の先頭から  $(h + 1)$  文字目から文字列  $\tilde{w}_j$  が開始する。

なお、以後の説明文中において、 $\tilde{w}_j$  を  $w_c$  と表記することもある。

## 【0 1 2 7】

追加

【0 1 2 8】

【数 2 7】

$$\cdot \tilde{w}_j = (\tilde{w}_{j1}, \tilde{w}_{j2}, \dots, \tilde{w}_{jL_j}, \tilde{w}'_{j0}, \tilde{w}'_{j1}, \tilde{w}'_{j2}, \dots, \tilde{w}'_{jL_j-1}, \tilde{w}'_{jL_j})$$

$L_j$  : 文字列  $\tilde{w}_j$  の文字数

$\tilde{w}_{jk}$  : 文字列  $\tilde{w}_j$  の  $k$  番目の文字  $\tilde{w}_{jk} \in C$

$\tilde{w}'_{jk}$  : 文字列  $\tilde{w}_j$  の  $k$  文字目と  $(k + 1)$  文字目の間の単語切れ目の有無

$\tilde{w}'_{jk} \in S, S = \{s_0, s_1, s_2\}$

$s_0$  : 切れ目である

$s_1$  : 切れ目でない

( $s_2$  : 行先頭 or 最後尾)

$$\tilde{w}'_{j0} = \tilde{w}'_{jL_j} = s_0$$

( $s_2$  は行先頭、最後尾を同一形式で表すために用意したものであり、本質的なものではない)

【0 1 2 9】

変更

$$\cdot \text{特徴 } r = (r_c, r_s)$$

$r_c$  : 文字特徴、 $r_s$  : 文字間特徴

追加

【0 1 3 0】

【数 2 8】

$$\cdot \text{文字特徴 } r_c = (r_{c1}, r_{c2}, r_{c3}, \dots, r_{cL})$$

$r_{ci}$  :  $i$  文字目の文字特徴 (= 文字認識結果)

(例. 1 位候補、3 位までの候補、類似度一定値以上の候補、

1, 2 位候補とその類似度、等)

$$\cdot \text{文字間特徴 } r_s = (r_{s0}, r_{s1}, r_{s2}, \dots, r_{sL})$$

$r_{si}$  :  $i$  文字目と  $(i + 1)$  文字目の間の文字間特徴

【0 1 3 1】

このとき、事後確率  $P(k_i | r)$  は以下の式で表わされる。



【0132】

【数29】

$$\begin{aligned}
 P(k_i | r) &= P(k_i | r_C, r_S) \\
 &= \frac{P(r_C, r_S | k_i) P(k_i)}{P(r_C, r_S)} \\
 &= \frac{P(r_C | r_S, k_i) P(r_S | k_i) P(k_i)}{P(r_C, r_S)}
 \end{aligned}
 \tag{23}$$

【0133】

ここで、 $P(r_S | k_i)$  と  $P(r_C | k_i)$  が独立であると仮定すると（これは文字特徴抽出と文字間特徴抽出が独立であることを意味する）、 $P(r_C | r_S, k_i) = P(r_C | k_i)$  であるから、上記式（23）は以下のようになる。

【0134】

【数30】

$$P(k_i | r) = \frac{P(r_C | k_i) P(r_S | k_i) P(k_i)}{P(r_C, r_S)} \tag{24}$$

【0135】

$P(r_C | k_i)$  は、前記式（13）とほぼ同様である。

【0136】

【数31】

$$\begin{aligned}
 P(r_C | k_i) &= P(r_{C1}, r_{C2}, \dots, r_{Ch} | k_i) \\
 &\quad \cdot P(r_{Ch+1} | \tilde{w}_{j1}) P(r_{Ch+2} | \tilde{w}_{j2}) \dots P(r_{Ch+L_j} | \tilde{w}_{jL_j}) \\
 &\quad \cdot P(r_{Ch+L_j+1}, \dots, r_{CL} | k_i) \\
 &= P(r_{C1}, r_{C2}, \dots, r_{Ch} | k_i) \left\{ \prod_{k=1}^{L_j} P(r_{Ch+k} | \tilde{w}_{jk}) \right\} \\
 &\quad \cdot P(r_{Ch+L_j+1}, \dots, r_{CL} | k_i)
 \end{aligned}$$

式 (25)

【0137】

$P(rs | k_i)$  は次のように表わされる。

【0138】

【数32】

$$\begin{aligned}
 P(r_s | k_i) &= P(r_{s1}, r_{s2}, \dots, r_{sh-1} | k_i) \\
 &\quad \cdot P(r_{sh} | \tilde{w}'_{j0}) P(r_{sh+1} | \tilde{w}'_{j1}) \cdots P(r_{sh+L_j} | \tilde{w}'_{jL_j}) \\
 &\quad \cdot P(r_{sh+L_j+1}, \dots, r_{sL-1} | k_i) \\
 &= P(r_{s1}, r_{s2}, \dots, r_{sh-1} | k_i) \left\{ \prod_{k=0}^{L_j} P(r_{sh+k} | \tilde{w}'_{jk}) \right\} \\
 &\quad \cdot P(r_{sh+L_j+1}, \dots, r_{sL-1} | k_i)
 \end{aligned}$$

式 (26)

【0139】

$P(k_i)$  は、4 節までと同様に求まっているものとする。ただし、一般に  $n(K)$  は4 節のときよりも大幅に増えることに注意する。

【0140】

## 5. 2. 実用に向けた近似

### 5. 2. 1. 文字列のない部分に対する近似と文字数正規化

4. 2. 1 節と同様の近似を前記式 (25) に用いると次のようになる。

【0141】

【数33】

$$P(r_c | k_i) = \prod_{k=1}^{L_j} P(r_{ch+k} | \tilde{w}_{jk}) \prod_{\substack{1 \leq k \leq h \\ h+L_j+1 \leq k \leq L}} P(r_{ck}) \quad \text{式 (27)}$$

【0142】

同様に、前記式 (26) は次のように近似される。

【0143】

【数34】

$$P(r_S | k_i) = \prod_{k=0}^{L_j} P(r_{Sh+k} | \tilde{w}_{jk}) \prod_{\substack{1 \leq k \leq h-1 \\ h+L_j+1 \leq k \leq L-1}} P(r_{Sk}) \quad \text{式 (28)}$$

【0144】

4. 2. 1節と同様、 $P(k_i | r) / P(k_i)$  という値を考えると、以下のように変形される。

【0145】

【数35】

$$\begin{aligned} \frac{P(k_i | r)}{P(k_i)} &= \frac{P(r_C | k_i) P(r_S | k_i)}{P(r_C, r_S)} \\ &\approx \frac{P(r_C | k_i)}{P(r_C)} \frac{P(r_S | k_i)}{P(r_S)} \quad \text{式 (29)} \\ &= \frac{P(k_i | r_C)}{P(k_i)} \frac{P(k_i | r_S)}{P(k_i)} \end{aligned}$$

【0146】

上記式(29)の1行目は、前記式(24)による。2行目は、

【0147】

【数36】

$$P(r_C, r_S) \approx P(r_C) P(r_S)$$

【0148】

という近似を用いている。上記式(29)は、「 $k_i$ である確率の、『特徴』を知ることによる変化」は $r_C$ と $r_S$ で独立に扱えることを示している。以下にそれぞれを計算する。

【0149】

【数37】

$$\begin{aligned} \frac{P(k_i | r_C)}{P(k_i)} &= \frac{P(r_C | k_i)}{P(r_C)} \\ &\approx \frac{\prod_{k=1}^{L_j} P(r_{Ch+k} | \tilde{w}_{jk}) \prod_{\substack{1 \leq k \leq h \\ h+L_j+1 \leq k \leq L}} P(r_{Ck})}{\prod_{k=1}^L P(r_{Ck})} \\ &= \prod_{k=1}^{L_j} \frac{P(r_{Ch+k} | \tilde{w}_{jk})}{P(r_{Ch+k})} \end{aligned} \quad \text{式 (30)}$$

【0150】

【数38】

$$\begin{aligned} \frac{P(k_i | r_S)}{P(k_i)} &= \frac{P(r_S | k_i)}{P(r_S)} \\ &\approx \frac{\prod_{k=0}^{L_j} P(r_{Sh+k} | \tilde{w}'_{jk}) \prod_{\substack{1 \leq k \leq h-1 \\ h+L_j+1 \leq k \leq L-1}} P(r_{Sk})}{\prod_{k=1}^{L-1} P(r_{Sk})} \\ &= \prod_{k=0}^{L_j} \frac{P(r_{Sh+k} | \tilde{w}'_{jk})}{P(r_{Sh+k})} \end{aligned} \quad \text{式 (31)}$$

【0151】

上記式(30)、式(31)の2行目の分母で用いている近似は、前記式(14)と同様である。なお、式(31)の3行目において、 $r_{s0}$ 、 $r_{sL}$ は必ず行先頭、最後尾《次節5.2.2の例ではd3》であるので、 $P(r_{s0}) = P(r_{sL}) = 1$ である。

【0152】

以上より、

【0153】

【数39】

$$\frac{P(k_i|r)}{P(k_i)} = \prod_{k=1}^{L_j} \frac{P(r_{Ch+k}|\tilde{w}_{jk})}{P(r_{Ch+k})} \prod_{k=0}^{L_j} \frac{P(r_{Sh+k}|\tilde{w}'_{jk})}{P(r_{Sh+k})} \quad \text{式 (32)}$$

【0154】

前記式(16)と同様、上記式(32)も文字列 $w_c$ の当てはまらない部分に関する記述はない。すなわち、この場合も「分母による正規化」の考え方ができる。

【0155】

### 5. 2. 2. 文字間特徴 $rs$ の一例

例として次のように定義する。

【0156】

・文字間特徴集合 $D = \{d_0, d_1, d_2, (, d_3)\}$

$d_0$  : 文字間が広い

$d_1$  : 文字間が狭い

$d_2$  : 接触している

( $d_3$  : 行の先頭または最後尾であり、必ず単語切れ目である)

・ $rs \in D$

このとき、

【0157】

【数40】

$$P(d_k | s_1)_{k=0,1,2} \quad l=0,1$$

【0158】

をあらかじめ求めておくことで、前記式(32)の第2項分子

【0159】

【数41】

$$P(r_{Sh+k}|\tilde{w}'_{jk})$$

【0160】

を求めることができる。ただし、 $P(d_3 | s_2) = 1$ である。

【0161】

また、

【0162】

【数42】

$$P(d_k)_{k=0,1,2}$$

【0163】

を求めておくことで、前記式(32)の第2項分母 $P(r_{sk})$ を求めることができる。

【0164】

### 5. 3. 具体例

4. 3節と同様に、英語の郵便物の宛名読取りにおいて、都市名を読むことを考える。

【0165】

図8は、都市名が書かれていると考えられる部分であり、全部で5文字であることがわかる。1-2文字目は接触しており、2-3文字目の間は広く、3-4文字目の間、4-5文字目の間は狭い。図9は単語辞書の内容であり、全ての都市名が格納されている。本例では、都市名は3候補存在する。各都市名の下に記載された記号( $s_0$ ,  $s_1$ )は単語切れ目であり、 $s_0$ は単語切れ目、 $s_1$ は単語切れ目でないところ、である。

【0166】

図10は、カテゴリの集合を図示したものである。カテゴリには位置情報が含まれるため、単語辞書とは異なる。カテゴリ $k_1$ は図9(a)の単語から作られ、カテゴリ $k_2$ ,  $k_3$ は図9(b)の単語から作られ、カテゴリ $k_4$ ,  $k_5$ ,  $k_6$ は図9(c)の単語から作られている。

【0167】

図8の各文字を文字認識にかけ、その文字認識結果が図10の各カテゴリの事後確率の計算に用いられる。計算に用いる特徴(=文字認識結果)は様々であるが、ここでは1位候補の文字を用いた例を示す。

【0168】

文字間の特徴も様々なものが考えられるが、ここでは5. 2. 2節で述べた例を用いる。図12は各文字間の特徴を示している。1-2文字目の間は接触しているため、文字間の特徴はd2である。2-3文字目の間は広いため、文字間の特徴はd0である。3-4文字目の間、4-5文字目の間は狭いため、文字間の特徴はd1である。

【0169】

5. 2. 1節で述べた近似を用いると、カテゴリk1の生起確率の、図11の文字認識結果を知ることによる変化 $P(k1 | rc) / P(k1)$ は、前記式(30)により、

【0170】

【数43】

$$\frac{P(k_1 | rc)}{P(k_1)} \approx \frac{P("S" | "S")}{P("S")} \frac{P("S" | "T")}{P("S")} \quad \text{式 (33)}$$

$$\cdot \frac{P("L" | "L")}{P("L")} \frac{P("I" | "I")}{P("I")} \frac{P("M" | "N")}{P("M")}$$

【0171】

である。カテゴリk1の生起確率の、図12の文字間特徴を知ることによる変化 $P(k1 | rs) / P(k1)$ は、前記式(31)により、

【0172】

【数44】

$$\frac{P(k_1 | rs)}{P(k_1)} \approx \frac{P(d_2 | s_1)}{P(d_2)} \frac{P(d_0 | s_0)}{P(d_0)} \frac{P(d_1 | s_1)}{P(d_1)} \frac{P(d_1 | s_1)}{P(d_1)} \quad \text{式 (34)}$$

【0173】

である。

【0174】

前記式(33)の計算をするために、3. 2. 2節、4. 2. 2節で述べた近似を用いれば、たとえば、 $p = 0.5$ 、 $n(E) = 26$ のときは、 $q = 0.02$ であるから、前記式(33)は次のように計算される。

【0175】

【数45】

$$\frac{P(k_1|z_C)}{P(k_1)} \approx p \cdot q \cdot p \cdot p \cdot q \cdot n(E)^5 \approx 594 \quad \text{式 (35)}$$

【0176】

前記式(34)の計算をするために、あらかじめ

【0177】

【数46】

$$P(d_k|s_1)_{k=0,1,2} \text{ と } P(d_k)_{k=0,1,2}$$

【0178】

を求めておく必要がある。例として、下記表1、表2の値が得られたとする。

【0179】

【表1】

$P(d_k, s_\ell)$  の値

$\ell \backslash k$	0 : 広い	1 : 狭い	2 : 接触	計
0 : 単語切れ目	$P(d_0, s_0)$ 0.16	$P(d_1, s_0)$ 0.03	$P(d_2, s_0)$ 0.01	$P(s_0)$ 0.2
1 : 単語切れ目でない	$P(d_0, s_1)$ 0.04	$P(d_1, s_1)$ 0.40	$P(d_2, s_1)$ 0.36	$P(s_1)$ 0.8
計	$P(d_0)$ 0.20	$P(d_1)$ 0.43	$P(d_2)$ 0.37	1

【0180】



【表2】

$P(d_k   s_\ell)$ の値			
$\ell \backslash k$	0 : 広い	1 : 狭い	2 : 接触
0 : 単語切れ目	$P(d_0   s_0)$ 0.8	$P(d_1   s_0)$ 0.15	$P(d_2   s_0)$ 0.05
1 : 単語切れ目 でない	$P(d_0   s_1)$ 0.05	$P(d_1   s_1)$ 0.50	$P(d_2   s_1)$ 0.45

【0181】

上記表1は

【0182】

【数47】

$$P(d_k \cap s_1)$$

【0183】

の値、表2は $P(d_k | s_1)$ の値を表にしたものである。この場合、

【0184】

【数48】

$$P(d_k \cap s_1) = P(d_k | s_1) P(s_1)$$

【0185】

の関係にあることに注意する。実際に前記式(34)の計算に必要なのは、 $P(d_k | s_1) / P(d_k)$ であるので、それを計算したものが下記表3である。

【0186】

【表 3】

$P(d_k   s_\ell) / P(d_k)$ の値			
$\ell \backslash k$	0 : 広い	1 : 狭い	2 : 接触
0 : 単語切れ目	$P(d_0   s_0)$ 4	$P(d_1   s_0)$ 0.35	$P(d_2   s_0)$ 0.14
1 : 単語切れ目 でない	$P(d_0   s_1)$ 0.25	$P(d_1   s_1)$ 1.16	$P(d_2   s_1)$ 1.22

【0187】

上記表3の値を基にして、前記式(34)は次のように計算される。

【0188】

【数49】

$$\frac{P(k_1 | r_s)}{P(k_1)} \approx 1.22 \cdot 4 \cdot 1.16 \cdot 1.16 \approx 6.57 \quad \text{式 (36)}$$

【0189】

前記式(29)より、カテゴリ  $k_1$  の生起確率の、図11の文字認識結果および図12の文字間特徴を知ることによる変化  $P(k_1 | r) / P(k_1)$  は、上記式(35)、式(36)の積で表わされるので、次のようになる。

【0190】

【数50】

$$\frac{P(k_1 | r)}{P(k_1)} \approx 594 \cdot 6.57 \approx 3900 \quad \text{式 (37)}$$

【0191】

同様に、 $k_2 \sim k_6$  についても、 $P(k_i | r_c) / P(k_i)$ 、 $P(k_i | r_s) / P(k_i)$ 、 $P(k_i | r) / P(k_i)$  を求めると、以下のようになる。

【0192】

【数51】

$$\begin{aligned}
\frac{P(k_2|r_C)}{P(k_2)} &\approx p \cdot q \cdot q \cdot q \cdot n(E)^4 \approx 1.83 \\
\frac{P(k_3|r_C)}{P(k_3)} &\approx p \cdot p \cdot p \cdot p \cdot n(E)^4 \approx 28600 \\
\frac{P(k_4|r_C)}{P(k_4)} &\approx p \cdot q \cdot q \cdot n(E)^3 \approx 3.52 \\
\frac{P(k_5|r_C)}{P(k_5)} &\approx p \cdot q \cdot q \cdot n(E)^3 \approx 3.52 \\
\frac{P(k_6|r_C)}{P(k_6)} &\approx q \cdot p \cdot p \cdot n(E)^3 \approx 87.9
\end{aligned}
\tag{38}$$

【0193】

【数52】

$$\begin{aligned}
\frac{P(k_2|r_S)}{P(k_2)} &\approx 1.22 \cdot 0.25 \cdot 1.16 \cdot 0.35 \approx 0.124 \\
\frac{P(k_3|r_S)}{P(k_3)} &\approx 0.14 \cdot 0.25 \cdot 1.16 \cdot 1.16 \approx 0.0471 \\
\frac{P(k_4|r_S)}{P(k_4)} &\approx 1.22 \cdot 0.25 \cdot 0.35 \approx 0.107 \\
\frac{P(k_5|r_S)}{P(k_5)} &\approx 0.14 \cdot 0.25 \cdot 1.16 \cdot 0.35 \approx 0.0142 \\
\frac{P(k_6|r_S)}{P(k_6)} &\approx 4 \cdot 1.16 \cdot 1.16 \approx 5.38
\end{aligned}
\tag{39}$$

【0194】

【数 5 3】

$$\begin{aligned}
 \frac{P(k_2|x)}{P(k_2)} &\approx 1.83 \cdot 0.124 \approx 0.227 \\
 \frac{P(k_3|x)}{P(k_3)} &\approx 28600 \cdot 0.0471 \approx 1350 \\
 \frac{P(k_4|x)}{P(k_4)} &\approx 3.52 \cdot 0.107 \approx 0.377 \\
 \frac{P(k_5|x)}{P(k_5)} &\approx 3.52 \cdot 0.0142 \approx 0.0500 \\
 \frac{P(k_6|x)}{P(k_6)} &\approx 87.9 \cdot 5.38 \approx 473
 \end{aligned}
 \tag{40}$$

【0 1 9 5】

上記式 (3 7)、式 (4 0) の中で最大なのはカテゴリ k1 である。したがって、推定結果は「ST LIN」となる。

【0 1 9 6】

文字間の特徴を用いない 4 節の方法では、前記式 (3 5)、式 (3 8) の中で最大であるカテゴリ k3 が推定結果となるが、文字間の特徴まで統合することで、総合的に最もマッチしていると考えられるカテゴリ k1 が選ばれていることがわかる。

【0 1 9 7】

このように、第 2 の実施の形態では、認識すべき単語に対応した入力文字列を各文字ごとに切出し、この文字切出しによって文字間の特徴を抽出するとともに、上記文字切出しによって得られる各文字に対してそれぞれ認識処理を行ない、あらかじめ認識すべき単語と単語の文字間の特徴の候補が格納されている単語辞書内の各単語の各文字および文字間の特徴を条件として文字認識結果として得られた特徴が出現する確率を求め、この求めた確率を文字認識結果として得られた特徴が出現する確率で除算し、単語辞書内の各単語の各文字および文字間の特徴に対してそれぞれ求めた上記各演算結果を全ての文字および文字間について乗算し、この乗算結果に基づき上記単語の認識結果を得るものである。

【0 1 9 8】

すなわち、文字認識結果を用いた単語認識において、少なくとも単語切出しの曖昧さをも考慮した、事後確率を基にした評価関数を用いることにより、単語切出しが確実でない場合にも精度良く単語認識が行なえる。

【0199】

次に、本発明の第3の実施の形態としての、文字接触、ノイズ混入のあるときのベイズ推定について説明する。この場合のベイズ推定は、文字接触、ノイズ混入が無視できないときに有用となる。

【0200】

#### 6. 文字接触、ノイズ混入の統合

5節までで述べた方法では、文字が必ず正しく切出されているという前提があり、これが満たされずに文字接触があるに対応できない。また、ノイズの混入にも対応できない。そこで、本節では、カテゴリを変更することにより、文字接触、ノイズの混入に対応したベイズ推定を行なう。

【0201】

#### 6. 1. 安定化

5節の定義を基に、次のように定義を追加変更する。

【0202】

変更

・カテゴリ  $K = \{k_i\}$

【0203】

【数54】

$k_i = (w_{jk}, h), w_{jk} \in W, W$ : 派生文字列の集合

なお、以後の説明文中において、 $w_{jk}$  を  $wd$  と表記することもある。

【0204】

追加

【0205】

【数 5 5】

・派生文字列

$$w_{jk} = (w_{jk1}, w_{jk2}, \dots, w_{jkL_{jk}}, w'_{jk0}, w'_{jk1}, \dots, w'_{jkL_{jk}})$$

$L_{jk}$  : 派生文字列  $w_{jk}$  の文字数

$w_{jk1}$  :  $w_{jk}$  の 1 番目の文字  $w_{jk1} \in C$

$w'_{jk1}$  :  $w_{jk}$  の 1 文字目と  $(1 + 1)$  文字目の間の単語切れ目の有無

$$w'_{jk1} \in S, \quad w'_{jk0} = w'_{jkL_{jk}} = s_0$$

・派生文字列  $w_{jk}$  と文字列  $\tilde{w}_j$  の関係

文字列  $\tilde{w}_j$  の 1 文字目と  $(1 + 1)$  文字目の間に作用  $a_{jk1} \in A$  が働くことで派生文字列  $w_{jk}$  ができるものとする。

$$A = \{a_0, a_1, a_2\} \quad a_0 : \text{作用なし}, \quad a_1 : \text{接触}, \quad a_2 : \text{ノイズ混入}$$

【0 2 0 6】

・  $a_0$  : 作用なし

該当 2 文字間に何もしない。

・  $a_1$  : 接触

該当 2 文字を接触させる。この作用により、該当 2 文字は 1 つの非文字に変換される。

例 「ONTARIO」の「T」と「A」とを接触  $\rightarrow$  ON#RIO (#は接触による非文字)

・  $a_2$  : ノイズ混入

該当 2 文字間にノイズ (非文字) を挿入する。

例 「ONT」の「N」と「T」との間にノイズを挿入  $\rightarrow$  ON\*T (\*はノイズによる非文字)

ただし、 $l = 0$ ,  $L_j$  のときは、それぞれ文字列  $w_c$  の左端、右端にノイズがあるとする。また、本定義では、ノイズが 2 文字以上連続して混入することはないと仮定している。

・ 非文字  $r \in C$

文字接触、ノイズ混入を考慮することによる非文字を同一視して  $r$  とし、文字集合  $C$  に含める。

【0207】

このとき、事後確率  $P(k_i | r)$  は前記式 (23)、式 (24) と同様である。

【0208】

【数56】

$$P(k_i | r) = \frac{P(r_C | k_i) P(r_S | k_i) P(k_i)}{P(r_C, r_S)} \quad \text{式 (41)}$$

【0209】

$P(r_C | k_i)$  もほぼ前記式 (25) と同様である。

【0210】

【数57】

$$P(r_C | k_i) = P(r_{C1}, r_{C2}, \dots, r_{Ch} | k_i) \left\{ \prod_{l=1}^{L_{jk}} P(r_{Ch+l} | w_{jkl}) \right\} \\ \cdot P(r_{Ch+L_{jk}+1}, \dots, r_{CL} | k_i) \quad \text{式 (42)}$$

【0211】

$P(r_S | k_i)$  もほぼ前記式 (26) と同様である。

【0212】

【数58】

$$P(r_S | k_i) = P(r_{S1}, r_{S2}, \dots, r_{Sh-1} | k_i) \left\{ \prod_{l=0}^{L_{jk}} P(r_{Sh+l} | w'_{jkl}) \right\} \\ \cdot P(r_{Sh+L_{jk}+1}, \dots, r_{SL-1} | k_i) \quad \text{式 (43)}$$

【0213】

6. 2.  $P(k_i)$  について

$P(w_c)$  がわかっているものとする。ここで、 $P(w_c)$  は、実際には、たとえば、郵便物の宛名読取りであれば、書状内での位置、行内での位置などの影

響を受けるが、それらの期待値として与えられるものとする。このとき、 $P(w_d)$  と  $P(w_c)$  は次のような関係にあると考えられる。

【0214】

【数59】

$$P(w_{jk}) = P(\tilde{w}_j) \left\{ \prod_{l=1}^{L_j-1} P(a_{jkl}) \right\} P(a_{jk0}) P(a_{jkl_j}) \quad \text{式 (44)}$$

【0215】

すなわち、文字接触とノイズ混入は、接触確率  $P(a_1)$ 、ノイズ混入確率  $P(a_2)$  を与えることで、5節までの枠組みに統合できる。ただし、上記式(44)で

【0216】

【数60】

$$P(a_{jk0}), P(a_{jkl_j})$$

【0217】

は両端のノイズの有無に関する項であり、一般に文字間と両端ではノイズの存在する確率が異なるため、ノイズ混入確率  $P(a_2)$  とは別に値を定めるものとする。

【0218】

$P(w_c)$  と  $P(w_c, h)$  や、 $P(w_d)$  と  $P(w_d, h)$  の関係は、先ほども述べたような影響（書状内での位置など）をどうモデル化し、近似するかによるため、ここでは触れない。

【0219】

### 6. 3. 非文字 $\gamma$ について

文字特徴として、3. 2. 1節のように、「1位候補の文字」を用いた場合を考える。非文字  $\gamma$  を特徴抽出したとき、出現する1位候補の文字はどれも同様に確からしいと考えられる。そこで、次のように非文字を扱う。

【0220】



【数61】

$$P(e_i|\gamma) = \frac{1}{n(E)} \quad \text{式 (45)}$$

【0221】

## 6. 4. 具体例

5. 3節と同様に、英語の郵便物の宛名読取りにおいて、都市名を読むことを考える。

【0222】

本節の特徴を分かりやすくするため、単語切出しが完全に成功しており、また、カテゴリに複数単語からなる文字列が存在しないという前提を設ける。図13は、都市名が書かれていると考えられる部分であり、5文字であることがわかる。図14は、単語辞書の内容であり、全ての都市名が格納されている。本例では、都市名は3つの候補がある。

【0223】

図15は、カテゴリ集合を図示したものであり、単語辞書を基に作られる派生文字列のうち、5文字であるものが列挙されている。全ての5文字の派生文字列を列挙すると、たとえば、「PETAR」より派生する「P#A\*R」なども含まなければならないが、6. 2節に述べた接触確率 $P(a1)$ 、ノイズ混入確率 $P(a2)$ がある程度以上小さい場合は無視することができる。本例では無視することにする。

【0224】

図13の各文字を文字認識にかけ、その文字認識結果を用いて図15の各カテゴリの事後確率の計算を行なう。計算に用いる特徴(=文字認識結果)は様々であるが、ここでは1位候補の文字を用いた例を示す。

【0225】

図16は、各文字の認識結果を示している。たとえば、一番左の文字は、認識結果の1位が「S」であったことを示す。カテゴリ $k2$  (S#TAL)の生起確率の、図16の文字認識結果を知ることによる変化 $P(k2|r)/P(k2)$ は、前記式(16)より、

【0226】

【数62】

$$\frac{P(k_2|r)}{P(k_2)} \approx \frac{P("S"|"S")}{P("S")} \frac{P("E"|"E")}{P("E")} \frac{P("T"|"T")}{P("T")} \frac{P("A"|"A")}{P("A")} \frac{P("L"|"L")}{P("L")} \quad \text{式 (46)}$$

【0227】

である。さらに、3. 2節、4. 2. 2節で述べた近似を用いれば、たとえば、 $p = 0.5$ 、 $n(E) = 26$ のときは、 $q = 0.02$ であるから、上記式(46)は次のように計算される。

【0228】

【数63】

$$\begin{aligned} \frac{P(k_2|r)}{P(k_2)} &\approx p \cdot \frac{1}{n(E)} \cdot p \cdot p \cdot p \cdot n(E)^5 \\ &= p \cdot p \cdot p \cdot p \cdot n(E)^4 \approx 28600 \end{aligned} \quad \text{式 (47)}$$

【0229】

計算過程を見ると、結局、非文字以外の4文字について計算していることと等価になっている。同様に、他のカテゴリについても計算される。ここでは代表として、大きな値を示すと容易に推察される  $k_6$ 、 $k_7$ 、 $k_8$  について計算する。

【0230】

【数64】

$$\begin{aligned}
\frac{P(k_6|r)}{P(k_6)} &\approx q \cdot p \cdot p \cdot p \cdot q \cdot n(E)^5 \approx 594 \\
\frac{P(k_7|r)}{P(k_7)} &\approx \frac{1}{n(E)} \cdot q \cdot p \cdot p \cdot p \cdot n(E)^5 \\
&= q \cdot p \cdot p \cdot p \cdot n(E)^4 \approx 1140 \quad \text{式 (48)} \\
\frac{P(k_8|r)}{P(k_8)} &\approx p \cdot \frac{1}{n(E)} \cdot p \cdot p \cdot p \cdot n(E)^5 \\
&= p \cdot p \cdot p \cdot p \cdot n(E)^4 \approx 28600
\end{aligned}$$

【0231】

これらの比較に際し、5節までは $P(k_i)$ を等しいと見なしていたが、本節では6.2節で述べたように、文字接触、ノイズ混入を考慮することによって、 $P(k_i)$ に変化が生じている。そこで、変化が生じる前の $P(k_i)$ を、全て等しいと見なして $P(k_i) = P_0$ とおくことにする。 $P_0$ は、前記式(44)における $P(w_c)$ であると考えることができる。また、変化が生じた後の $P(k_i)$ は、前記式(44)における $P(w_d)$ と考えられる。したがって、変化が生じた後の $P(k_i)$ は次のようになる。

【0232】

【数65】

$$P(k_i) = P_0 \left\{ \prod_{l=1}^{L_j-1} P(a_{jk_l}) \right\} P(a_{jk_0}) P(a_{jk_{L_j}}) \quad \text{式 (49)}$$

【0233】

ここで、接触確率 $P(a_1) = 0.05$ 、文字間ノイズ混入確率 $P(a_2) = 0.002$ 、両端ノイズ混入確率 $P'(a_2) = 0.06$ とすると、たとえば、 $P(k_2)$ は次のように計算される。

【0234】

【数 6 6】

$$P(k_2) = P_0 \cdot 0.948 \cdot 0.05 \cdot 0.948 \cdot 0.948 \cdot 0.948 \cdot 0.94 \cdot 0.94 \\ \approx 0.0357 P_0$$

式 (50)

【0 2 3 5】

計算中、文字間において文字接触もノイズ混入もない確率  $P(a_0) = 1 - P(a_1) - P(a_2) = 0.948$ 、両端においてノイズ混入のない確率  $P'(a_0) = 1 - P'(a_2) = 0.94$  を用いている。

【0 2 3 6】

同様にして、 $P(k_6)$ 、 $P(k_7)$ 、 $P(k_8)$  を計算すると、

【0 2 3 7】

【数 6 7】

$$P(k_6) = P_0 \cdot 0.948 \cdot 0.948 \cdot 0.948 \cdot 0.948 \cdot 0.94 \cdot 0.94 \\ \approx 0.714 P_0$$

$$P(k_7) = P_0 \cdot 0.948 \cdot 0.948 \cdot 0.948 \cdot 0.06 \cdot 0.94 \\ \approx 0.0481 P_0$$

$$P(k_8) = P_0 \cdot 0.002 \cdot 0.948 \cdot 0.948 \cdot 0.94 \cdot 0.94 \\ \approx 0.00159 P_0$$

式 (51)

【0 2 3 8】

上記式 (50)、式 (51) を前記式 (47)、式 (48) に用いて変形すると、以下ようになる。

【0 2 3 9】

【数 6 8】

$$P(k_2|r) \approx 28600 \cdot 0.0357 P_0 \approx 1020 P_0$$

$$P(k_6|r) \approx 594 \cdot 0.714 P_0 \approx 424 P_0$$

$$P(k_7|r) \approx 1140 \cdot 0.0481 P_0 \approx 54.8 P_0$$

$$P(k_8|r) \approx 28600 \cdot 0.00159 P_0 \approx 45.5 P_0$$

式 (52)

【0240】

参考までに他のカテゴリについて同様な計算をすると、以下のようになる。

【0241】

【数69】

$$P(k_1|x) \approx 40.7 P_0, P(k_3|x) \approx 40.7 P_0,$$

$$P(k_4|x) \approx 1.63 P_0, P(k_5|x) \approx 0.0653 P_0,$$

$$P(k_9|x) \approx 1.81 P_0, P(k_{10}|x) \approx 0.0727 P_0,$$

$$P(k_{11}|x) \approx 0.0880 P_0$$

【0242】

以上より、事後確率が最も大きいのはカテゴリ  $k_2$  であり、図13に書かれている都市名は「SISTAL」で、「I」と「S」とが接触していると推定される。

【0243】

このように、第3の実施の形態は、単語辞書内の各単語の各文字には、文字に加えて非文字の情報を含ませ、かつ、非文字の情報を含む文字の単語の出現確率は、非文字の情報を含まない文字の単語の出現確率を基に設定することにより、文字認識結果を用いた単語認識において、文字接触やノイズ混入をも考慮した、事後確率を基にした評価関数を用いて単語認識を行なうことができ、これにより、文字接触やノイズ混入がある場合にも精度良く単語認識が行なえる。

【0244】

次に、本発明の第4の実施の形態としての、文字の切出しが一意に定まらないときのベイズ推定について説明する。この場合のベイズ推定は、漢字や仮名など、分離のある文字では有用である。また、英語における筆記体など、実際の文字どうしの切れ目以外にも多くの切れ目の候補を出さざるを得ない場合にも有用である。

【0245】

## 7. 文字切出しの統合

6節までで述べてきた方法では、文字自体が分離することはないという前提が

あった。しかし、漢字や仮名など、文字自体が2つ以上に分離する場合もある。たとえば、『明』という漢字は、文字切出しを行なうと、へんとつくりが別々に文字候補として切出される。このとき、2つの文字候補を結合すべきか、別々にすべきか、により、複数の文字切出し候補が現われる。

## 【0246】

このような文字の分離には、6節までの方法では対応できない。また、逆に文字接触が多く、それを切る処理をしている場合、実際に接触しているところだけでなく、文字自体を切ってしまうこともある。後で詳しく述べるが、認識の戦略として文字自体の切断をある程度許容する方が、性能がよい。この場合も同様に、6節までの方法では対応できない。そこで、本節では、文字の分離による複数の文字切出し候補に対応したベイズ推定を行なう。

## 【0247】

## 7. 1. 文字切出しについて

文字が接触することのある対象における文字切出しにおいては、文字接触を切る処理が行なわれる。この処理において、「文字の切れ目でない箇所」が切れ目候補に挙がる場合と、「文字の切れ目」が切れ目候補に挙がらない場合を比べると、後者の方が認識に悪影響を与える。理由は次の通りである。

## 【0248】

- ・「文字の切れ目でない箇所」が切れ目候補に挙がる場合

「切れ目候補で切った場合」と「切らない場合」を試すことができるので、切り過ぎたことで正しい文字切出しがなくなるわけではない。

- ・「文字の切れ目」が切れ目候補に挙がらない場合

正しい文字切出しを得る手段はない。

## 【0249】

このため、文字切出しにおいて、文字の切れ目以外にも多くの切れ目候補を挙げておくことは有用である。しかし、「切れ目候補で切った場合」と「切らない場合」を試すということは、複数の文字切出しパターン候補があることになる。6節までで述べてきた方法では、異なる文字切出しパターン候補間の比較はできない。そこで、以下に説明する方法でこれを解決する。

【0250】

## 7. 2. 定式化

6節の定義を基に、次のように定義を追加変更する。

【0251】

### 変更

- ・切れ目状態の集合  $S = \{s_0, s_1, s_2, s_3\}$   
 $s_0$  : 単語切れ目である、 $s_1$  : 文字切れ目である、 $s_2$  : 文字切れ目ではない、( $s_3$  : 行先頭または最後尾)  
 5節以降で定義していた『切れ目である』は「単語切れ目である」の意味であり、 $s_0$  にあたる。『切れ目でない』は  $s_1$  と  $s_2$  に分かれたことになる。
- ・  $L$  : 切れ目候補で分割された部分 (これを「セル」と呼ぶことにする) の個数

### 追加

【0252】

【数70】

- ・ ユニット  $u_{ij} (i \leq j)$   
 $i$  番目のセルから  $(j-1)$  番目のセルまで結合したもの

【0253】

### 変更

- ・ カテゴリ  $K = \{k_i\}$

【0254】

【数71】

- $k_i = (w_{jk}, m_{jk}, h), w_{jk} \in W$   
 $m_{jk} = (m_{jk1}, m_{jk2}, \dots, m_{jkL_{jk}}, m_{jkL_{jk}+1})$   
 $m_{jk1}$  : 文字  $w_{jk1}$  の当てはまるユニットの先頭セル番号。ユニットは  $u_{m_{jk1}m_{jk1}+1}$  と表される。  
 $h$  : 派生文字列  $w_{jk}$  の位置。  $(h+1)$  番目のセルから派生文字列  $w_{jk}$  が開始する。

【0255】

追加

【0256】

【数72】

・切れ目パターン  $k'_i = (k'_{i0}, k'_{i1}, \dots, k'_{iL_C})$

$k'_i$  :  $k_i$  における切れ目状態

$L_C$  : 派生文字列  $w_{jk}$  のあてはまっている全ユニットに含まれる全セル数

$L_C = m_{jkL_{jk}+1} - m_{jk1}$

$k'_{i1}$  :  $(h+1)$  番目のセルと  $(h+1+1)$  番目のセルの間における切れ

目の状態  $k'_{i1} \in S$

$$k'_{i1} = \begin{cases} s_0 & (\text{単語切れ目のとき、つまり、} \\ & \exists n, w'_{jkn} = s_0, 1 = m_{jkn+1} - h - 1 \text{ のとき}) \\ s_2 & (\forall n, 1 \neq m_{jkn} - h - 1 \text{ のとき}) \\ s_1 & (\text{上記以外のとき}) \end{cases}$$

【0257】

変更

【0258】

【数73】

・文字特徴

$r_C = (r_{C12}, r_{C13}, r_{C14}, \dots, r_{C1L+1}, r_{C23}, r_{C24},$   
 $\dots, r_{C2L+1}, \dots, r_{CLL+1})$

$r_{Cn_1n_2}$  : ユニット  $u_{n_1n_2}$  の文字特徴

・文字間特徴  $r_S = (r_{S0}, r_{S1}, \dots, r_{SL})$

$r_{Sn}$  :  $n$  番目のセルと  $(n+1)$  番目のセルの間の文字間特徴

【0259】

このとき、事後確率  $P(k_i | r)$  は、やはり前記式 (23)、式 (24) と同様である。

【0260】



【数 7 4】

$$P(k_i | r) = \frac{P(r_C | k_i) P(r_S | k_i) P(k_i)}{P(r_C, r_S)} \quad \text{式 (53)}$$

【0 2 6 1】

$P(r_C | k_i)$  は次のように表わされる。

【0 2 6 2】

【数 7 5】

$$\begin{aligned} P(r_C | k_i) &= P(r_{Cm_{jk1}m_{jk2}} | w_{jk1}) P(r_{Cm_{jk2}m_{jk3}} | w_{jk2}) \cdots \\ &\quad \cdot P(r_{Cm_{jkL_{jk}}m_{jkL_{jk}+1}} | w_{jkL_{jk}}) \\ &\quad \cdot P(\cdots, r_{n_1n_2}, \cdots | k_i) \\ &\quad \text{それ以外の } (n_1n_2) \\ &= \left\{ \prod_{n=1}^{L_{jk}} P(r_{Cm_{jkn}m_{jkn+1}} | w_{jkn}) \right\} \\ &\quad \cdot \left\{ \begin{array}{c} P(\cdots, r_{n_1n_2}, \cdots | k_i) \\ n_1, n_2 \\ \forall b, 1 \leq b \leq L_{jk}, (n_1, n_2) \neq (m_{jkb}, m_{jkb+1}) \end{array} \right\} \end{aligned} \quad \text{式 (54)}$$

【0 2 6 3】

$P(r_S | k_i)$  は次のように表わされる。

【0 2 6 4】

【数 7 6】

$$\begin{aligned} P(r_S | k_i) &= P(r_{S1}, r_{S2}, \cdots, r_{Sh-1} | k_i) \\ &\quad \cdot P(r_{Sh} | k'_{i0}) P(r_{Sh+1} | k'_{i1}) \cdots P(r_{Sh+L_C} | k'_{iL_C}) \\ &\quad \cdot P(r_{Sh+L_C+1}, \cdots, r_{SL-1} | k_i) \end{aligned} \quad \text{式 (55)}$$

【0 2 6 5】

$P(k_i)$  は、本節ではカテゴリ  $k_i$  に  $m_{jk}$  が含まれているため、その影響を

考慮する必要がある。個々の文字の当てはまるユニットの形状、ユニットに当てはまっている文字、近傍ユニット間の形状バランスなどが影響すると考えられるが、そのモデル化についてはここでは触れない。

【0266】

### 7. 3. 実用に向けた近似

#### 7. 3. 1. 文字列のない部分に対する近似と文字数正規化

4. 2. 1 節と同様の近似を前記式 (54) に用いると、次のようになる。

【0267】

【数77】

$$P(r_C | k_i) \approx \prod_{n=1}^{L_{jk}} P(r_{Cn_{jkn}m_{jkn+1}} | w_{jkn})$$

$$\prod_{n_1, n_2} P(r_{Cn_1 n_2})$$

$$\forall b, 1 \leq b \leq L_{jk}, (n_1, n_2) \neq (m_{jkb}, m_{jkb+1})$$

式 (56)

【0268】

実際には、 $r_{cn_1n_3}$ と $r_{cn_1n_2}$ 、 $r_{cn_2n_3}$ との間には何らかの相関があると考えられるため、この近似は4. 2. 1 節のときよりもさらに粗いものである。

【0269】

また、前記式 (55) も同様に近似すると次のようになる。

【0270】

【数78】

$$P(r_S | k_i) = \prod_{n=0}^{L_C} P(r_{Sh+n} | k'_{in}) \prod_{\substack{1 \leq n \leq h-1 \\ h+L_C+1 \leq n \leq L-1}} P(r_{Sn})$$

式 (57)

【0271】

さらに、5. 2. 1 節と同様に、 $P(k_i | r) / P(k_i)$  を計算すると次のようになる。

【0272】

【数79】

$$\frac{P(k_i|x)}{P(k_i)} \approx \frac{P(k_i|x_C)}{P(k_i)} \frac{P(k_i|x_S)}{P(k_i)}$$

$$\approx \prod_{n=1}^{L_{jk}} \frac{P(x_{Cm_{jkn}m_{jkn+1}}|w_{jkn})}{P(x_{Cm_{jkn}m_{jkn+1}})} \prod_{n=0}^{L_C} \frac{P(x_{Sh+n}|k'_{in})}{P(x_{Sh+1})}$$

式 (58)

【0273】

前記式(32)と同様、上記式(58)も派生文字列 $w_d$ の当てはまっていない部分に関する記述はなく、「分母による正規化」の考え方ができる。

【0274】

7. 3. 2. 切れ目と文字間特徴 $r_s$ について

6節までと違い、本節では切れ目状態として $s_2$ （文字切れ目でない）があるので、文字間特徴集合として5. 2. 2節と同様に $D$ を用いた場合、

【0275】

【数80】

$$P(d_k | s_1)_{k=0,1,2} \quad l=0,1,2$$

【0276】

を求めておけばよいことになる。ここで注意しなければならないのは、これらは全て7. 1節で述べたように、「切れ目候補」として挙げた部分における話であることである。 $s_2$ （文字切れ目でない）は、「切れ目候補として挙げたが切れ目ではない」という意味であり、

【0277】

【数81】

$$P(d_k | s_2)_{k=0,1,2}$$

【0278】

を求める際に気をつける必要がある。

【0279】

【数82】

$P(d_k)_{k=0,1,2}$

【0280】

を求める際も同様である。

【0281】

#### 7. 4. 具体例

6. 4 節と同様に、英語の郵便物の宛名読取りにおいて、都市名を読むことを考える。

【0282】

本節の特徴を分かりやすくするため、単語切出しが完全に成功しており、カテゴリに複数の単語からなる文字列が存在せず、ノイズの混入もなく、文字切出しによって全ての文字切れ目が検出されている（すなわち、6 節のようにノイズ、接触文字に関するカテゴリが必要ない）とする。

【0283】

図17は、都市名が書かれていると考えられる部分であり、5つのセルがある。図18は、考えられる文字切出しパターン候補である。本例では、簡単のため、セル2と3、セル4と5の間は必ず切れることがあらかじめ分かっているとした（切れない確率がごく小さいため、無視したと考えてもよい）。

【0284】

すると、切れ目候補はセル1と2、セル3と4の間であり、考えられる文字切出しパターン候補を列挙すると、図18のようになる。図19は単語辞書の内容であり、全ての都市名が格納されている。本例では、都市名は3候補ある。

【0285】

図20は、カテゴリ集合を図示したものである。単語切出しが完全に成功しているという前提なので、「BAYGE」は図18(a)、「RAGE」は図18(b)と(c)、「ROE」は図18(d)にのみ当てはまる。

【0286】

図18に現れる各ユニットを文字認識にかけ、その文字認識結果が図20の力

カテゴリの事後確率の計算に用いられる。計算に用いる特徴(=文字認識結果)は様々であるが、ここでは1位候補の文字を用いた例を示す。

【0287】

図21は、各ユニットの認識結果を示している。たとえば、セル1と2とを結合したユニットは、認識結果の1位が「R」であったことを示す。

【0288】

文字間特徴も様々なものが考えられるが、ここでは5.2.2節で述べた例を簡略化して、次のようなものを用いる。

【0289】

・文字間特徴集合  $D' = \{d'1, d'2\}$

$d'1$  : 接触していない

$d'2$  : 接触している

図22は、セル1-2間、3-4間の文字間特徴を示している。1-2間は接触しておらず、3-4間は接触している。

【0290】

7.3.1節で述べた近似を用いると、カテゴリ  $k_i$  (BAYGE) の生起確率の、図21の認識結果を知ることによる変化  $P(k_i | rc) / P(k_1)$  は、前記式(58)より、

【0291】

【数83】

$$\frac{P(k_1 | rc)}{P(k_1)} \approx \frac{P("B" | "B")}{P("B")} \frac{P("A" | "A")}{P("A")} \frac{P("A" | "Y")}{P("A")} \frac{P("G" | "G")}{P("G")} \frac{P("E" | "E")}{P("E")} \quad \text{式(59)}$$

【0292】

である。カテゴリ  $k_i$  の生起確率の、図22の文字間特徴を知ることによる変化  $P(k_i | rs) / P(k_1)$  は、前記式(58)より、

【0293】

【数 8 4】

$$\frac{P(k_1|r_s)}{P(k_1)} \approx \frac{P(d'_1|s_1)}{P(d'_1)} \frac{P(d'_2|s_1)}{P(d'_2)} \quad \text{式 (60)}$$

【0 2 9 4】

である。

【0 2 9 5】

前記式 (59) の計算をするために、3. 2. 2 節、4. 2. 2 節で述べた近似を用いれば、たとえば、 $p = 0.5$ 、 $n(E) = 26$  のときは、 $q = 0.02$  であるから、前記式 (59) は次のように計算される。

【0 2 9 6】

【数 8 5】

$$\frac{P(k_1|r_c)}{P(k_1)} \approx p \cdot p \cdot q \cdot p \cdot p \cdot n(E)^5 \approx 14900 \quad \text{式 (61)}$$

【0 2 9 7】

前記式 (60) の計算をするために、あらかじめ

【0 2 9 8】

【数 8 6】

$$P(d'_k|s_1)_{k=1,2} \quad 1=1,2 \text{ と } P(d'_k)_{k=1,2}$$

【0 2 9 9】

を求めておく必要がある。例として下記表 4、表 5 の値が得られたとする。

【0 3 0 0】

【表4】

$P(d_k', s_\ell)$  の値

$\ell \backslash k$	1 : 非接触	2 : 接触	計
1 : 文字切れ目	$P(d_1', s_1)$ 0.45	$P(d_2', s_1)$ 0.05	$P(s_1)$ 0.5
2 : 文字切れ目 でない	$P(d_1', s_2)$ 0.01	$P(d_2', s_2)$ 0.49	$P(s_2)$ 0.5
計	$P(d_1')$ 0.46	$P(d_2')$ 0.54	1

【0301】

【表5】

$P(d_k' | s_\ell)$  の値

$\ell \backslash k$	1 : 非接触	2 : 接触
1 : 文字切れ目	$P(d_1'   s_0)$ 0.90	$P(d_2'   s_1)$ 0.10
2 : 文字切れ目 でない	$P(d_1'   s_2)$ 0.02	$P(d_2'   s_2)$ 0.98

【0302】

上記表4は

【0303】

【数87】

$$P(d_k \cap s_1)$$

【0304】

の値、表5は $P(d_k | s_1)$  の値を表にしたものである。この場合、

【0305】

【数88】

$$P(d_k \cap s_1) = P(d_k | s_1) P(s_1)$$

【0306】

の関係にあることに注意する。実際に前記式(60)の計算に必要なのは、 $P(d_k | s_1) P(d_k)$  であるので、それを計算したものが下記表6である。

【0307】

【表6】

$P(d_{k'}   s_\ell) / P(d_{k'})$ の値		
$\ell \backslash k$	1 : 非接触	2 : 接触
1 : 文字切れ目	1.96	0.19
2 : 文字切れ目でない	0.043	1.18

【0308】

上記表6の値を基にして、前記式(60)は次のように計算される。

【0309】

【数89】

$$\frac{P(k_1 | r_s)}{P(k_1)} \approx 1.96 \cdot 0.19 \approx 0.372 \quad \text{式(62)}$$

【0310】

前記式(60)より、カテゴリ  $k_1$  の生起確率の、図21の文字認識結果および図22の文字間特徴を知ることによる変化  $P(k_1 | r) / P(k_1)$  は、前記式(61)、式(62)の積で表わされるので、次のようになる。

【0311】

【数90】

$$\frac{P(k_1 | r)}{P(k_1)} \approx 14900 \cdot 0.19 \approx 2830 \quad \text{式(63)}$$

【0312】

同様に、 $k_2 \sim k_4$  についても、 $P(k_i | r_c) / P(k_i)$ 、 $P(k_i | r_s) / P(k_i)$ 、 $P(k_i | r) / P(k_i)$  を求めると、以下のようにな



る。

【0313】

【数91】

$$\begin{aligned}\frac{P(k_2|r_C)}{P(k_2)} &\approx q \cdot p \cdot q \cdot p \cdot n(E)^4 \approx 45.7 \\ \frac{P(k_3|r_C)}{P(k_3)} &\approx p \cdot p \cdot p \cdot p \cdot n(E)^4 \approx 28600 \\ \frac{P(k_4|r_C)}{P(k_4)} &\approx p \cdot p \cdot p \cdot n(E)^3 \approx 2197\end{aligned}\quad \text{式 (64)}$$

【0314】

【数92】

$$\begin{aligned}\frac{P(k_2|r_S)}{P(k_2)} &\approx 1.96 \cdot 1.81 \approx 3.55 \\ \frac{P(k_3|r_S)}{P(k_3)} &\approx 0.043 \cdot 0.19 \approx 0.00817 \\ \frac{P(k_4|r_S)}{P(k_4)} &\approx 0.043 \cdot 1.81 \approx 0.0778\end{aligned}\quad \text{式 (65)}$$

【0315】

【数93】

$$\begin{aligned}\frac{P(k_2|r)}{P(k_2)} &\approx 45.7 \cdot 3.55 \approx 162 \\ \frac{P(k_3|r)}{P(k_3)} &\approx 28600 \cdot 0.00817 \approx 249 \\ \frac{P(k_4|r)}{P(k_4)} &\approx 2197 \cdot 0.0778 \approx 171\end{aligned}\quad \text{式 (66)}$$

【0316】

これらの比較に際し、5節までは $P(k_i)$ を等しいと見なしていたが、本節では文字の形状を考慮に入れる。

【0317】

図 1 8 (d) は各ユニットの幅が揃っている。図 1 8 (a) はそれに次いで揃っている。しかし、図 1 8 (b)、(c) は不揃いである。

【0 3 1 8】

この揃っているかどうかの度合いを何らかの方法でモデル化して  $P(k_i)$  に反映させれば、より正確な単語認識が可能になるが、ここではその方法は問わない。

【0 3 1 9】

本例では、その結果として、

【0 3 2 0】

【数 9 4】

$$P(k_1) : P(k_2) : P(k_3) : P(k_4) = 2 : 1 : 1 : 10 \quad \text{式 (67)}$$

【0 3 2 1】

になったとする。比例定数を  $P_1$  とし、前記式 (6 3)、式 (6 6) に上記式 (6 7) を用いて変形すると、以下ようになる。

【0 3 2 2】

【数 9 5】

$$\begin{aligned} P(k_1|r) &\approx 2830 \cdot 2P_1 \approx 5660P_1 \\ P(k_2|r) &\approx 162 \cdot P_1 \approx 162P_1 \\ P(k_3|r) &\approx 249 \cdot P_1 \approx 249P_1 \\ P(k_4|r) &\approx 171 \cdot 10P_0 \approx 1710P_0 \end{aligned} \quad \text{式 (68)}$$

【0 3 2 3】

以上より、事後確率が最も大きいのはカテゴリ  $k_1$  であり、図 1 7 に書かれている都市名は「BAYGE」であると推定される。

【0 3 2 4】

図 2 1 の文字認識による結果が最も大きいのは、前記式 (6 1)、式 (6 4) よりカテゴリ  $k_3$  であり、図 2 2 の文字間特徴による結果が最も大きいのは、前記式 (6 2)、式 (6 5) よりカテゴリ  $k_2$  であり、文字形状のバランスの評価が最も高いのはカテゴリ  $k_4$  であるが、全てを統合した推定を行なうことで、カ

テゴリ k1 を選ぶことができる。

【 0 3 2 5 】

このように、第 4 の実施の形態では、認識すべき単語に対応した入力文字列を各文字ごとに切出すとともに、この文字切出しによって文字間の接離を考慮した複数種類の切出し結果を求め、この求めた全ての切出し結果の各文字に対してそれぞれ認識処理を行ない、あらかじめ認識すべき単語と単語の文字間の特徴の候補が格納されている単語辞書内の各単語の各文字および文字間の特徴を条件として文字認識結果として得られた特徴が出現する確率を求め、この求めた確率を文字認識結果として得られた特徴が出現する確率で除算し、単語辞書内の各単語の各文字および文字間の特徴に対してそれぞれ求めた上記各演算結果を全ての文字および文字間について乗算し、この乗算結果に基づき上記単語の認識結果を得るものである。

【 0 3 2 6 】

すなわち、文字認識結果を用いた単語認識において、少なくとも文字切出しの曖昧さをも考慮した、事後確率を基にした評価関数を用いることにより、文字切出しが確実でない場合にも精度良く単語認識が行なえる。

【 0 3 2 7 】

【発明の効果】

以上詳述したように本発明によれば、単語の文字数が一定でない場合においても精度良く単語認識が行なえる単語認識方法および単語認識プログラムを記憶した記憶媒体を提供できる。

【 0 3 2 8 】

また、本発明によれば、単語切出しが確実でない場合にも精度良く単語認識が行なえる単語認識方法および単語認識プログラムを記憶した記憶媒体を提供できる。

【 0 3 2 9 】

また、本発明によれば、文字接触やノイズ混入がある場合にも精度良く単語認識が行なえる単語認識方法および単語認識プログラムを記憶した記憶媒体を提供できる。

【0330】

さらに、本発明によれば、文字切出しが確実でない場合にも精度良く単語認識が行なえる単語認識方法および単語認識プログラムを記憶した記憶媒体を提供できる。

【図面の簡単な説明】

【図1】

本発明の実施の形態に係る単語認識方法を実現するための単語認識システムの構成を概略的に示すブロック図。

【図2】

認識すべき単語の記載例を示す図。

【図3】

図2の記載例に対する単語辞書の内容を示す図。

【図4】

図2の記載例に対する文字認識結果を示す図。

【図5】

認識すべき単語の他の記載例を示す図。

【図6】

図5の記載例に対する単語辞書の内容を示す図。

【図7】

図5の記載例に対する文字認識結果を示す図。

【図8】

認識すべき単語のさらに他の記載例を示す図。

【図9】

図8の記載例に対する単語辞書の内容を示す図。

【図10】

図8の記載例に対するカテゴリの集合を示す図。

【図11】

図8の記載例に対する文字認識結果を示す図。

【図12】

図 8 の記載例に対する文字間の特徴を示す図。

【図 13】

認識すべき単語のさらに他の記載例を示す図。

【図 14】

図 13 の記載例に対する単語辞書の内容を示す図。

【図 15】

図 13 の記載例に対するカテゴリの集合を示す図。

【図 16】

図 13 の記載例に対する文字認識結果を示す図。

【図 17】

認識すべき単語のさらに他の記載例を示す図。

【図 18】

図 17 の記載例に対する文字切出しパターン候補を示す図。

【図 19】

図 17 の記載例に対する単語辞書の内容を示す図。

【図 20】

図 17 の記載例に対するカテゴリの集合を示す図。

【図 21】

図 18 の文字切出しパターン候補に対する各ユニットの認識結果を示す図。

【図 22】

図 17 の記載例に対する文字間特徴を示す図。

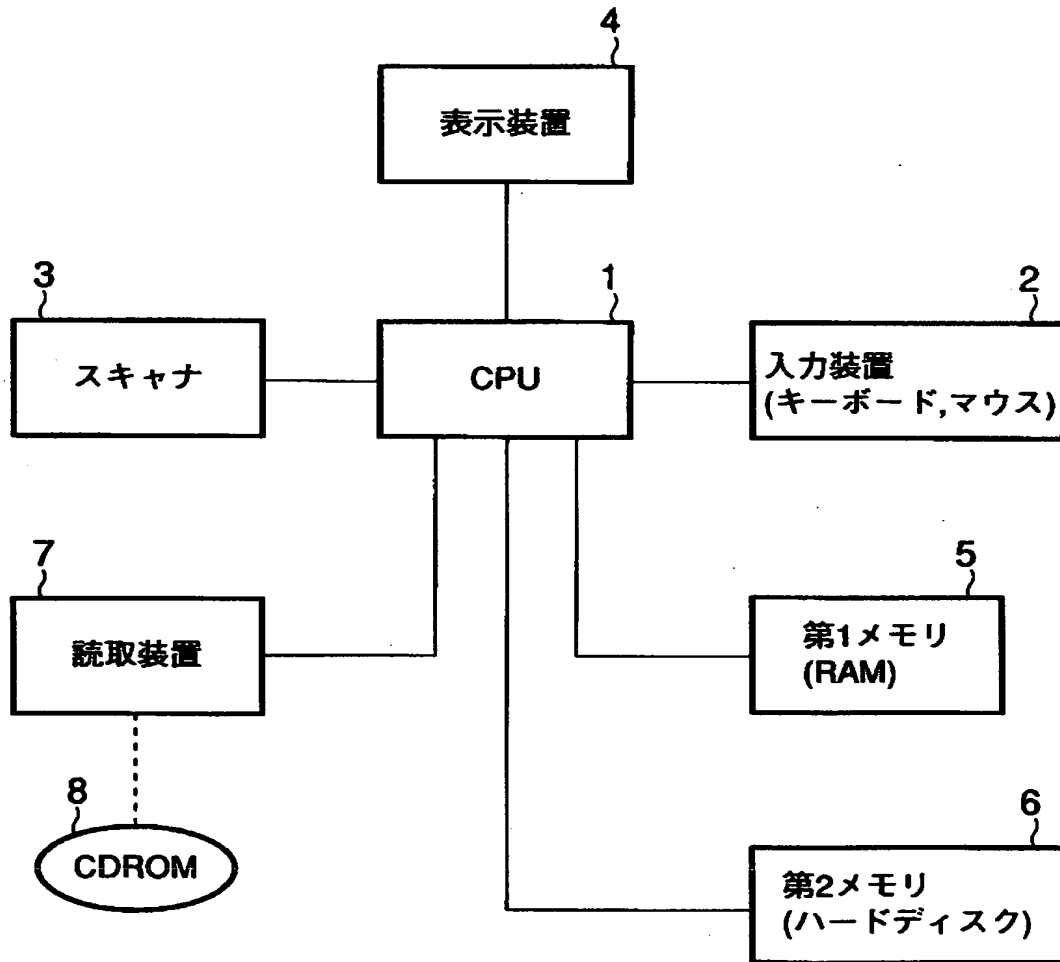
【符号の説明】

- 1 ……CPU
- 2 ……入力装置
- 3 ……スキャナ（画像入力手段）
- 4 ……表示装置
- 5 ……第 1 メモリ
- 6 ……第 2 メモリ
- 7 ……読取装置

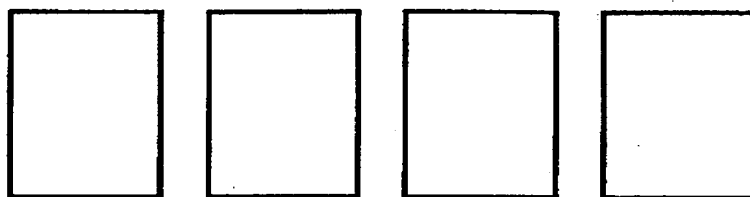
8 …… C D - R O M ( 記 憶 媒 体 )

【書類名】 図面

【図1】



【図 2】

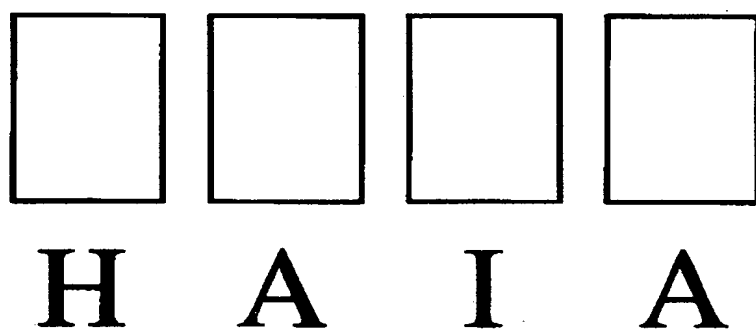


【図 3】

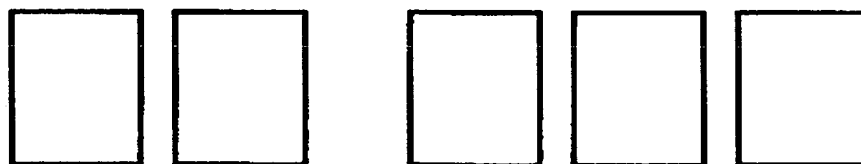
$k_1$	MAIR
$k_2$	SORD
$k_3$	ABLA
$k_4$	HAMA
$k_5$	HEWN



【図 4】



【図 5】



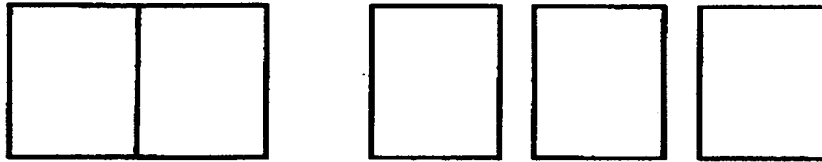
【図 6】

$k_1$	COH
$k_2$	LE ITH
$k_3$	OTH
$k_4$	SK
$k_5$	ST LIN

【図 7】

S	K	C	T	H

【図 8】



【図 9】

- (a)  $\begin{matrix} \tilde{w}_1 \\ \tilde{w}_1 \end{matrix} \begin{array}{|c|} \hline \text{ST LIN} \\ \hline s_0 \ s_1 \ s_0 \ s_1 \ s_1 \ s_0 \\ \hline \end{array}$
- (b)  $\begin{matrix} \tilde{w}_2 \\ \tilde{w}_2 \end{matrix} \begin{array}{|c|} \hline \text{SLIM} \\ \hline s_0 \ s_1 \ s_1 \ s_1 \ s_0 \\ \hline \end{array}$
- (c)  $\begin{matrix} \tilde{w}_3 \\ \tilde{w}_3 \end{matrix} \begin{array}{|c|} \hline \text{SIM} \\ \hline s_0 \ s_1 \ s_1 \ s_0 \\ \hline \end{array}$

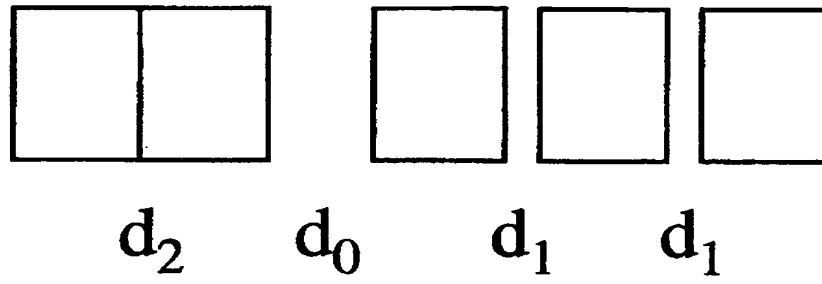
【図 1 0】

$k_1$	STLIN
$k_2$	SLIM
$k_3$	SLIM
$k_4$	SIM
$k_5$	SIM
$k_6$	SIM

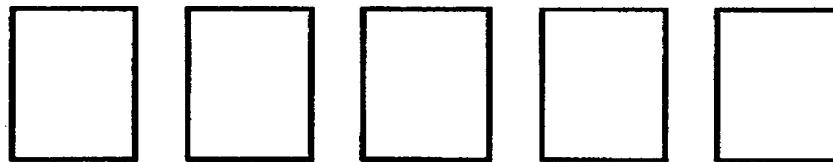
【図 1 1】

S	S	L	I	M

【図 1 2】



【図 1 3】



【図 1 4】

$\tilde{w}_1$	SISTAL
$\tilde{w}_2$	PETAR
$\tilde{w}_3$	STAL

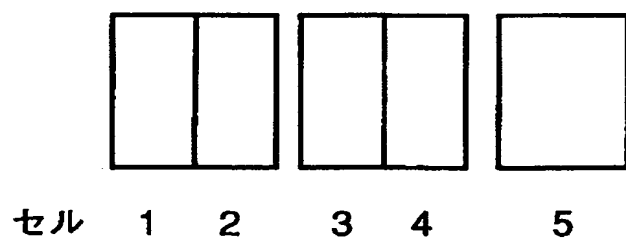
【図 15】

$K_1(W_{11})$	#STAL
$K_2(W_{12})$	S#TAL
$K_3(W_{13})$	SI#AL
$K_4(W_{14})$	SIS#L
$K_5(W_{15})$	SIST#
$K_6(W_{21})$	PETAR
$K_7(W_{31})$	*STAL
$K_8(W_{32})$	S*TAL
$K_9(W_{33})$	ST*AL
$K_{10}(W_{34})$	STA*L
$K_{11}(W_{35})$	STAL*

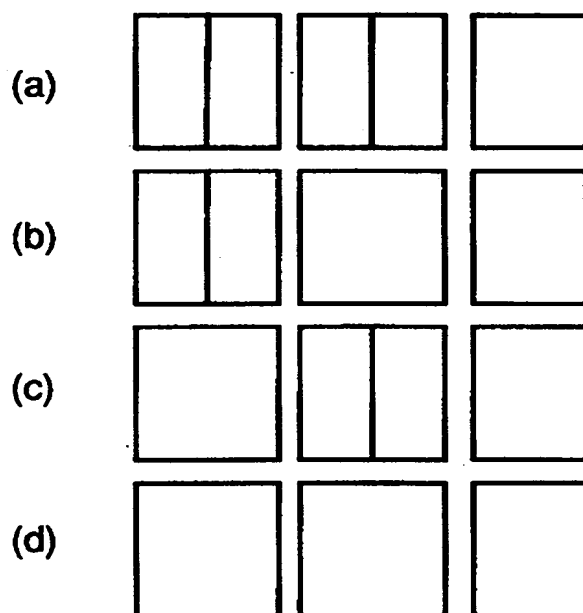
【図 16】

S	E	T	A	L

【図 1 7】



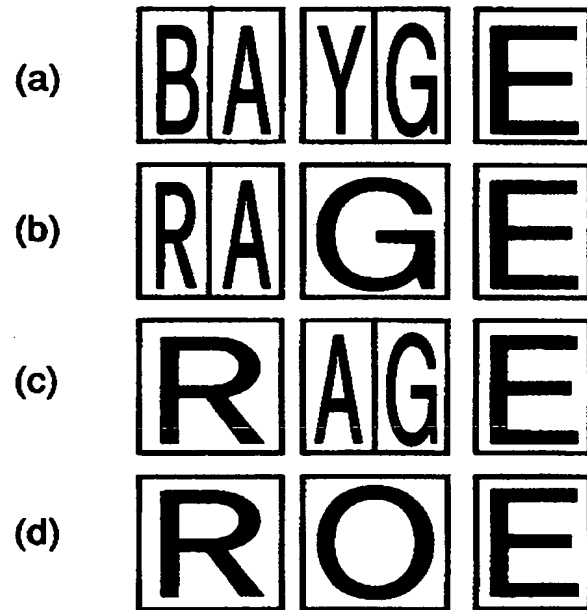
【図 1 8】



【図 1 9】

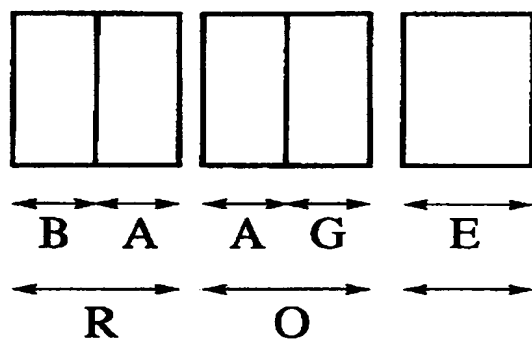


【図 2 0】

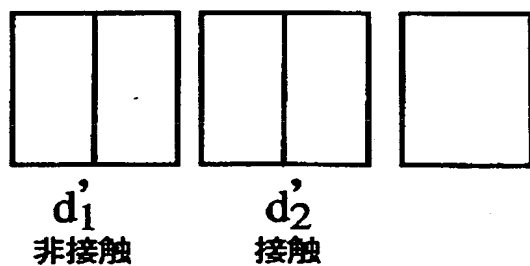




【图 2 1】



【图 2 2】



【書類名】 要約書

【要約】

【課題】 単語の文字数が一定でない場合においても精度良く単語認識が行なえる単語認識方法を提供する。

【解決手段】 文字認識結果を用いた単語認識において、認識すべき単語に対応した入力文字列を各文字ごとに認識処理を行ない、あらかじめ認識すべき単語の候補が格納されている単語辞書内の各単語の各文字を条件として文字認識結果として得られた特徴が出現する確率を求め、この求めた確率を文字認識結果として得られた特徴が出現する確率で除算し、単語辞書内の各単語の各文字に対してそれぞれ求めた上記各除算結果を全ての文字について乗算し、この乗算結果に基づき上記単語の認識結果を得る。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000003078]

1. 変更年月日	1990年 8月22日
[変更理由]	新規登録
住 所	神奈川県川崎市幸区堀川町72番地
氏 名	株式会社東芝